

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Ingegneria Civile, Chimica, Ambientale e dei Materiali

Ciclo XXXIII

**Settore Concorsuale:** 08/A4

**Settore Scientifico Disciplinare:** ICAR/06

3D HIGH-RESOLUTION TECHNIQUES APPLIED ON SMALL AND MEDIUM  
SIZE OBJECTS: FROM THE ANALYSIS OF THE PROCESS TOWARDS  
QUALITY ASSESSMENT

**Presentata da:** Chiara Francolini

**Coordinatore Dottorato**

Prof. Luca Vittuari

**Supervisore**

Prof. Gabriele Bitelli

**Esame finale anno 2021**

# Summary

<b>SUMMARY .....</b>	<b>i</b>
<b>ABSTRACT.....</b>	<b>iv</b>
<b>1. THE ACQUISITION OF THE 3D METRIC DATA.....</b>	<b>1</b>
<b>1.1 Computer Vision and Artificial Intelligence .....</b>	<b>1</b>
<b>1.2 Human Vision System (HVS) .....</b>	<b>2</b>
<b>1.3 Image processing.....</b>	<b>3</b>
<b>1.4 Sensing devices .....</b>	<b>4</b>
1.4.1 Shape acquisition: the 3D sensing devices.....	4
1.4.2 Optical passive and active sensors .....	6
<b>2. IMAGE BASED METHODS: CLOSE-RANGE PHOTOGRAMMETRY .....</b>	<b>9</b>
<b>2.1 Generalities: photogrammetric process .....</b>	<b>9</b>
<b>2.2 Image acquisition: digital cameras .....</b>	<b>10</b>
<b>2.3 Camera calibration.....</b>	<b>12</b>
<b>2.4 Image measurement .....</b>	<b>17</b>
2.4.1 Edges extraction.....	17
2.4.2 Image segmentation .....	18
2.4.3 Features extraction.....	19
<b>2.5 Image matching and 3D object reconstruction.....</b>	<b>20</b>
2.5.1 Feature-based matching and Structure from Motion approach (SfM).....	20
2.5.2 Area-based matching .....	22
2.5.3 3D object reconstruction .....	27
<b>2.6 Multi-image acquisition .....</b>	<b>31</b>
2.6.1 Geometric rules.....	31
2.6.2 Acquisition procedure: 3x3 rules .....	33

<b>2.7 Quality assessment of a close-range photogrammetric process: a case study</b>	<b>36</b>
2.7.1 Multi-view camera configuration .....	39
2.7.2 Photo alignment optimization .....	41
2.7.3 Results and considerations .....	43
<b>3. RANGE BASED METHODS: FROM LASER TO STRUCTURED-LIGHT TECHNOLOGY</b>	<b>65</b>
<b>3.1 History and developments</b> .....	<b>65</b>
<b>3.2 Measurement principles</b> .....	<b>65</b>
3.2.1 Triangulation systems .....	66
3.2.2 Structure-light systems (SLS) .....	69
3.2.3 Terrestrial Laser Scanners.....	72
<b>3.3 Active sensors calibration issues</b> .....	<b>79</b>
<b>3.4 Calibration of Structured-light scanners</b> .....	<b>81</b>
3.4.1 Calibration artifacts.....	82
3.4.2 Camera and projector models .....	83
3.4.3 Calibration procedure.....	85
<b>3.5 Artec Spider and Artec Eva scanners</b> .....	<b>87</b>
3.5.1 Calibration.....	89
<b>4. UNCONVENTIONAL APPLICATIONS OF HIGH-RESOLUTION ACQUISITION TECHNIQUES</b>	<b>92</b>
<b>4.1 Introduction</b> .....	<b>92</b>
<b>4.2 High-resolution acquisition techniques for Cultural Heritage</b> .....	<b>93</b>
<b>4.3 Longobard Basin of Santo Stefano Church, Bologna</b> .....	<b>94</b>
4.3.1 The restoration .....	95
4.3.2 The historic interpretation.....	97
<b>4.4 Clay tablet and seal of Ancarani Collection, Civic Archaeologic Museum, Bologna</b> .....	<b>114</b>
4.4.1 3D acquisition and 3D model reconstruction .....	115

4.4.2 Digital enhancement .....	116
4.4.3 Vectorization.....	119
<b>4.5 Safe of the Obello Firmo <i>domus</i> in Pompeii.....</b>	<b>124</b>
4.5.1 Photogrammetric survey: planning and images acquisition .....	126
4.5.2 Data processing and 3D model generation .....	127
4.5.3 3D reconstruction of the original appearance and shape .....	129
<b>5. CONCLUSIONS .....</b>	<b>130</b>
<b>Bibliography .....</b>	<b>132</b>

## ABSTRACT

The need for metric data acquisition is an issue strictly related to the human capability of describing the world with rigorous and repeatable methods. The first attempts of metric reproduction of a scene dates to many centuries ago, but it was only at the end of XIX century when the interest in 3D objects acquisition began to give satisfactory results. From the invention of photography to the development of advanced computers, the metric data acquisition has been subjected to rapid and impressive mutations involving both the 3D reconstruction methods and algorithms, and the acquisition devices. Nowadays, there exists a strict and solid connection between metric data acquisition and image processing, Computer Vision and Artificial Intelligence, that have led to the possibility to link the Geomatics studies with the Computer Vision field.

The sensor devices for the 3D model generation of an object are various and characterized by different functioning principles, varying the acquisition method and the applied technology. In this work, optical passive and active sensors are treated as recently developed techniques that in the last few decades have been the subject of studies in practical and academic fields. The interest falls on their versatility in terms of application fields, accuracy of final results and instrumentation technology developments. This work focuses specifically on close-range photogrammetry, Time of Flight (ToF) sensors and Structured-light scanners (SLS), as they represent three modern solutions for the acquisition of 3D metric data.

Within this thesis, starting from the functioning principles of the techniques and showing some issues related to them, the aim is to study the data produced by the mentioned high-resolution techniques considering small and medium size objects. The work will highlight their potentialities in terms of acquisition and resulting 3D digital models, analyzing the fundamental and most critical steps of the process leading to the quality assessment of the data. Central themes are the instruments calibration, the acquisition plan and the interpretation of the final results.

One of the crucial topics in the research is the capability of the acquisition techniques to satisfy unconventional requirements in several different fields. Specifically, the field of Cultural Heritage (CH) is considered, since in the last decades has shown great interest in close-range photogrammetry and SLS. The 3D models generated with these technologies, in fact, are versatile products suitable for restoration, cataloguing, enhancement and divulgation of the heritage. The studies involving CH and high-resolution acquisition are therefore characterized by a strong interdisciplinary connotation, that makes these themes to have a central role in both academic research and application fields.

The thesis starts with an overview about the history and developments of 3D metric data acquisition and underlines the correlation between the techniques for the three-dimensional modeling and the Computer Vision. Chapter 1 also treats the Human Vision System as a starting point to understand how sensing devices for shape acquisition have been developed starting from it. Then, a complete overview of 3D sensing devices and a focus on optical active and passive sensors follows. In the next two chapters, image-based and range-based acquisition techniques are investigated. Chapter 2 starts from the enunciation of the basic principle of close-range photogrammetry considering digital cameras functioning principles, calibration issues, and the process leading to the 3D mesh reconstruction. The case of multi-image acquisition is analyzed, deepening the quality assessment of the photogrammetric process through a case study. Chapter 3 is devoted to the range-based acquisition techniques, namely ToF laser scanners and SLSs. In this section, the two

technologies are studied and the functioning principles behind them are shown; moreover, issues related to both are considered. Specifically, the range accuracy of ToF laser scanner Riegl VZ400 and the calibration issue of Spider and Eva Artec structured-light projection scanners are studied. Lastly, Chapter 4 focuses on unconventional applications of the mentioned high-resolution acquisition techniques showing some examples of study cases in the field of Cultural Heritage on which innovative analysis methodologies are applied.

# 1. THE ACQUISITION OF THE 3D METRIC DATA

The history of the modern techniques for the metric data acquisition different from traditional topographic techniques dates back to the first developments in the field of photography at the end of the XIX century, when Albrecht Meydenbauer, a German architect, used photographs to measure the façades of buildings, through the direct measurement of orientation angles. At the same time, the invention of the stereoscope has permitted to deepen the concept of photogrammetric measurement: the new born stereo-comparators in fact, allowed the calculation and recording of points coordinates through the setting of measuring marks in two stereoscopic images, thus marking the beginning of the analogue photogrammetry era.

After few decades and due to the advent of electronic computers, the attention for the analogic photographic support, which limited the output accuracy, disappeared, turning into the will to use purely analytical approach. It was 1950s when the first photogrammetric analytical applications began to give impressive results in terms of precision and statistical information. From that moment on, the new developed bundle techniques for photogrammetric triangulation began to be applied to several photogrammetric applications, and software for airborne and close-range photogrammetry appeared.

It was the end of XX Century and the advantages provided by the bundle adjustment were such impressive that couldn't be ignored; furthermore, great improvements were done and several camera calibration methods began to be studied, especially for close-range applications, where bundle adjustment helped the development of self-calibration. Metric cameras could be replaced by any non-metric camera, thus making the whole process easier. Due to the development of image sensors and computer technology, nowadays we can talk about digital photogrammetry. This last step of photogrammetric science involves not only digital images taken with advanced digital cameras, but also digital procedures for the photogrammetric process.

The advent of digital techniques for the acquisition of metric data in fact, has directly involved the Computer Vision (CV) research field, that enables the processing of digital images with a high level of automation in several phases, from the automatic feature extraction to their measurement. During the first years of XXI Century, CV algorithms became even more effective, and thanks to the development of projective geometry, pattern recognition interest operators and Structure from Motion approach, they assume nowadays a central role in photogrammetric reconstruction.

## 1.1 Computer Vision and Artificial Intelligence

Computer Vision is a scientific discipline that studies how digital images or video sequences can be understood from computer to infer something about a given scene of the world: the goal is the production of numerical information describing a given sight, which means interpreting visual data with a visual language (Prince, 2012). In general, CV tasks are the acquisition, processing, analysis, and understanding of visual data. Specifically, CV provides the automated extraction of information form visual data of different nature, such as images, video sequences (bi-dimensional data), point cloud or meshes (three-dimensional data), and this operation includes the development of algorithms for numerical models involving geometry, physics, statistics as well as computational learning theory (Forsyth and Ponce, 2003). CV itself is a branch of a wider discipline, Artificial Intelligence (AI),

since its goals to use computer to emulate the human vision, including learning and acting based on visual inputs, that is, indeed, proper of the AI. For this reason, in order to understand the CV, it is necessary the knowledge of AI basic principles.

The basis for the AI theory enunciation were laid in the early 1940s by the American mathematician Claude Shannon. He is considered the founding father of the information theory: Shannon studied how any logical relationship between electrical signals could be governed by the Boolean algebra (1937), setting, de facto, the basis for the modern analysis and design of digital systems for information management. Few years later (1948-1949) he published two scientific papers about the cryptography theory, after being came in contact with Alan Turing, the leading British mathematician. Turing, after the contribution to national British defense as cryptanalyst during World War II, in those year was interested in neurology and physiology studies to understand the relationship between computer and nature. From the combined studies of these two personalities, and thanks also to two Harvard students, Minsky and Edmonds, for the creation the first neural network, around 1950s the AI theory was beginning to assume a more and more interesting and articulate shape. The expression Artificial Intelligence however, was introduced in 1956 by John McCarty during a workshop at Dartmouth College, attended by the best mathematicians of the period (including Shannon and Minsky). From that moment, the field of AI research started its brilliant progresses.

In the late 1960s in the academic world, CV could have been the first step for the application of Artificial Intelligence (AI) to robots: CV, in fact, infers information about the environment to be navigated by robotic systems using the learning theory and recognition technique. As AI wants to endow computers with information-processing capabilities as similar as possible to those proper of the biological organisms, so as CV was born to replicate the human vision structure and build a vision system capable to accomplish autonomously the tasks proper to the human system (Ballard and Brown, 1982; Huang, 1996). A common mathematical model used by CV (but also image and video processing) is the human vision system model (HVS) that aims to simplify the human visual system and formalize the biological and physiological processes that govern it.

## **1.2 Human Vision System (HVS)**

HVS is composed by the sensory organ (eyes), the connecting pathways through to the primary visual cortex and other parts of the central nervous system. The cornea and the lens, located in the eyes, behave as a mobile and flexible compound lens that refract light in the visible part of the spectrum (400-780 nm): about 60-70% of the light is well transmitted by the cornea-lens system, whilst the 20-30% is lost due to the scattering of the ocular media. Only the 10% is absorbed by the photoreceptor pigment. Cornea and lens accomplish the first step of the vision, namely the projection of the visual scene, and invert image onto the retina.

The cornea and the surfaces of the lens are well approximated by spherical surfaces with defined radii of curvature. Their refractive power is positive and the focal length is 22.6 mm. The active deformation of the lens (accommodation) causes the focal length variation from 53 mm to 32 mm: in optimal optical condition, the projected image of a distant object is in focus when the accommodation is relaxed. The retina is a light sensitive layer of tissue of the innermost part of the eye. Its task is to analyze the projected scene through the photoreceptor cells located at its backside, the rods and the cones. Although they are both responsible of the vision phototransduction, their functions are different and complementary: rods operate in poor visibility situations when there is a low level of

light and provide the so-called scotopic vision (also called “night vision”). Cones indeed, allow the so-called photopic vision: they enable the central vision (reading, driving, face recognition, etc.) since their response times to stimuli are faster. However, rods and cone effectively transduce the image into electrical pulses that reach the optic chiasm through the optic nerve, decussating and branching.

The primary visual cortex (V1) of the brain intercepts in its right and left half respectively, the information coming from the eyes: the scene acquired by the right field of view is processed from the left side of the brain and, vice versa, the left one from the right side. All the information travel through the optic tract to the lateral geniculate nucleus (LGN), which is the center of transmission and processing of the visual information, together with the V1, from which receives reciprocal innervation. Specifically, the primary visual cortex selects visual information that, flowing through the visual pathway, reach the secondary visual area (composed by V2, V3, V4, V5/MT sub areas) where visual primitives are processed (Hornberg, 2017). The human vision process described above is what the CV aims to mimic. This is possible thanks to the combination of devoted hardware and software: the algorithms that set the rules of the vision process aims to reproduce the steps of the human vision process and regulate the performances of vision systems, which are the hardware component. The vision systems, therefore, are those technologies governed by the theoretic principles of the CV discipline.

### **1.3 Image processing**

Computer Vision and Image Processing are two related scientific fields. As Gonzalez and Wood affirm (Gonzalez and Wood, 2008), there are no precise boundaries between them: both CV and Image Processing, in fact, deal with visual inputs taking advantage of algorithms to process them, and encompass varied and wide of fields of application. Moreover, CV itself is a branch of a wide discipline, Artificial Intelligence (AI), since it goals to use computer to emulate the human vision, including learning and acting based on visual inputs, that is, indeed, proper of the AI. However, it is common to make a distinction among low-, mid- and high-level computerized processes, to make order in this continuum of fields. The processes differ by input and output type: some low- level processes are, for example, the noise reduction and the image sharpening, where both input and output are digital images. On the other hand, a mid-level process is characterized by the fact that the outputs are, instead, attributes extracted from the images (input), as in the case of object classification and object extraction operations. Finally, the high-level processes are those that perform the cognitive functions associated with vision, i.e. those processes that understand the object detected in the previous level.

Image processing birth is strictly related to the advent of the modern digital computers and only in the early 1960s they become powerful enough to carry out the first significant image processing tasks. One of the first application of the digital technique dates to 1964, when images of the moon taken by the U.S spacecraft was enhanced and restored. From 1970s image processing began to be used in other field such as medical imaging, remote Earth observation and astronomy, and nowadays, since the field has grown strongly, it has successful application also in biology, defense, archaeology and industry.

## 1.4 Sensing devices

In general, sensing devices are sensors designed to be comparable with the human eyes. Computer vision algorithms that aims to enable the vision systems to convert the light intensity to an information, exploit the vision functions. Such as the array of photoreceptors located in the retina, so as in sensing devices a film or an electronic sensor, enable to convert the light in a signal.

The most common sensing device is a classical 2D camera, a projective system that register the visual content of an environment scene. Modern cameras provide bi-dimensional digital images representing the scene from a fixed point of view and that contain color information. The images are the basic input for the simplest form of CV, the 2D Computer Vision, that processes and analyses images or videos. Recently, both hardware and software has made huge headway and therefore has become possible to cope the depth estimation issue. This task is accomplished by the three-dimension (3D) computer vision: it takes advantage of innovative vision systems, which are capable to overcome the bi-dimensionality of the images and ensure a complete shape acquisition. The hardware components are 3D devices capable to capture range maps aside the color images and their employment is wide spread in numerous different fields (Giancola et. al, 2018).

### 1.4.1 Shape acquisition: the 3D sensing devices

The complete acquisition of the shape of a scene in three-dimensional form is possible thanks to 3D capturing techniques realized by 3D sensing devices. They are sensors capable to acquire the geometrical information of the environment and reconstruct its three-dimensional shape by capturing several scene points characterizing its geometrical features and building a triangulated network (that constitute a 3D accurate representation of the real scene. The methodologies for the shape acquisition are numerous (Sansoni et. Al, 2009; Giancola et. al, 2018; Guidi and Remondino, 2012): they exploit different tasks and are suitable for different application in a wide range of fields. Following the most general classification, the 3D object measurement and reconstruction technologies are categorized in contact and non-contact (Figure 1.1).

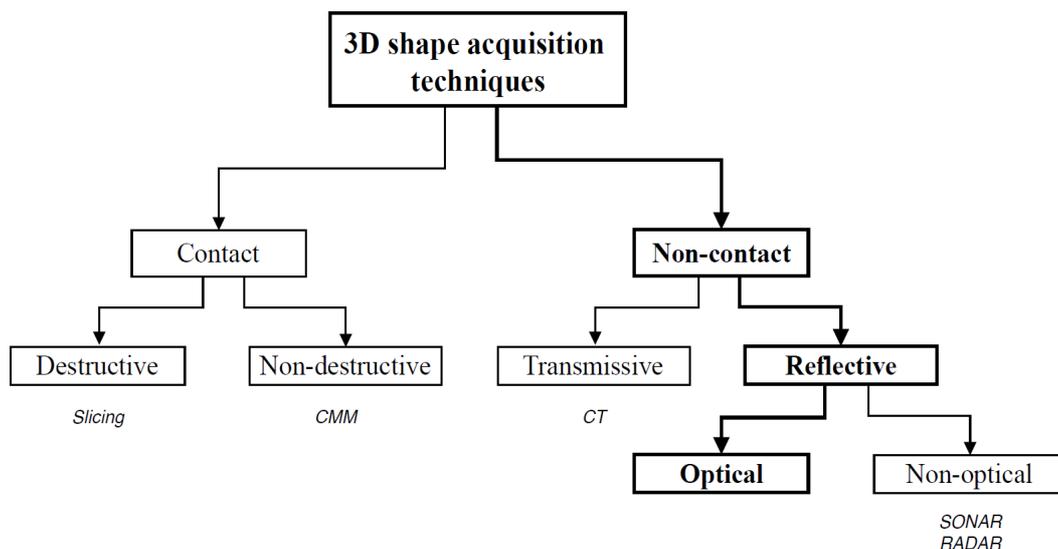


Figure 1.1. Taxonomy of 3D shape acquisition techniques.

The first category includes the techniques that require a contact between the object and the acquisition device. They can lead to the destruction of the object, and in that case, they are classified as destructive; or can preserve the original shape, and they are recognized as non-destructive. The first class of techniques (destructive) involves the slicing of the object with planar sections close to each other. 3D shape reconstruction takes advantage of the images of each slice: after the acquisition, the images are successively assembled together to create contours and internal structures of the object. Slicing technique is particularly invasive since the original shape of the object is destroyed and can no longer be restored.

The contact non-destructive acquisition techniques instead, use advanced devices capable of acquiring the coordinates of the object touching it in some peculiar points without any shape modification. This is the case of tactile measurement machines, or coordinate measuring machines (CMM): these commercial devices punctually touch the object with a measurement probe following a measurement path that covers the object and measures its geometry. If on the one hand, this system reaches values of accuracy in the order of microns, the contact can damage (but not destroy) both the object and the probe, phenomenon that is also influenced by the temperature of the object surface.

Nowadays, the 3D shape acquisition is mainly performed using non-contact techniques that are less invasive. These techniques, unlike the non-destructive CMM systems that accurately search for 3D points, usually acquire areas on the objects to be measured. For example, the Computer Axial Tomography (CAT), or just CT (Computer Tomography), is a well-known example of contactless transmissive technique. Invented in the early 1970s and largely used in medical diagnosis, it takes advantage of X-ray signal to identify the density changes within a body. The information is stored as a set of “slices” (images), which constitutes a 3D rendition of the inside of the body. On the contrary, contactless reflective techniques analyze the signal reflection. Depending on which wavelengths they use, we can make a distinction between optical and non-optical techniques. The optical techniques exploit the reflection of signal in the visible (400-800 nm) and Infra-Red (IR) (0.8-1000  $\mu\text{m}$ ) spectrum to get information about the environment. If the wavelengths analysed are not comprises within the mentioned ones, we refer to non-optical techniques, such as Sound Detection and Ranging (SONAR) and Radio Detection and Ranging (RADAR). They estimate respectively the time the sound and radio signals run through the environment to create range maps.

Nowadays, non-contact optical systems are those who are more employed for 3D model generation, using passive or active sensors. These techniques are those of which technological and functional aspects will be examined in depth in the next paragraphs.

### 1.4.2 Optical passive and active sensors

Among optical techniques, a first distinction exists between passive and active sensors (Figure 1.2).

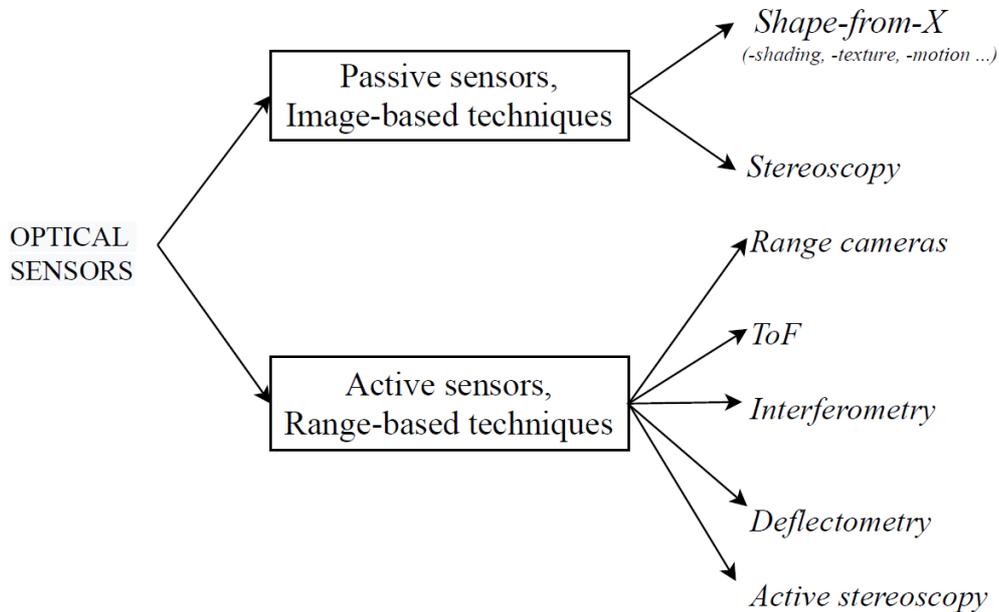


Figure 1.2. Optical sensors taxonomy.

Passive technologies measure the geometry of objects using the reflection of natural light on their surface from single or multi-stations. This method relies on observing a fixed scene and conveying it into an optical system able to extract geometric information from the view.

Nowadays the sensors are often very portable and low-cost. They are involved in the acquisition of the scene providing for 2D images from which the 3D object coordinates are calculated through mathematical models. This process is known as image-based modelling (IBM) and includes shape-from-X techniques. In general, shape-from-X techniques are those techniques that recover shape from image characteristics: X can stand for several characteristics, such as texture, specularity, contour, gradients, motion etc. A typical and largely used passive method is shape from shading, the CV process of computing the 3D shape of a surface using the light from only one image of that surface (Horn and Brooks, 1989; Zhang et Al. 1999). From this prospective, the classical stereo imaging is a shape-from-X method, and the characteristic to recover the 3D shape is the disparity: the apparent measurable motion of pixels between a pair of stereo images (Hornberg, 2017). Consequently, the photogrammetry is part of the image-based techniques for 3D shape reconstruction. A camera acquires a couple or a group of images of an object from different positions, and the system estimates the 3D coordinates of the scene from the correspondences between them.

Photogrammetry is used by a wide range of fields, from the traditional mapping and cartographic surveys to video games industry and 3D city modelling, from medical field to heritage documentation. Depending on the scale, photogrammetry is categorized as aerial photogrammetry (cartographic surveys) or close-range photogrammetry, used for small objects, such as the survey of

buildings or small mechanical components. During the last decades, digital photogrammetry technique developed significantly and now, it is not considered anymore a manual and time-consuming procedure. The improvement regards mostly the procedures, which nowadays are semi- or fully- automated and the technological progresses of the instrumentation used for the images acquisition, and next section on image-based methods will treat in depth the theme.

Range-based modelling relies on active sensors. This kind of sensors was developed essentially after 50s, when the interest in computer increased and laser technology was invented. Active sensors allow a direct acquisition of the 3D shape: they collect a set of three-dimensional coordinates referred to a very large number of points on the object surface and provide a highly detailed and accurate representation of most shapes. However, since the active sensors have an active role in point measurement, the accuracy and repeatability of 3D object coordinates is affected by the coding method and sensor calibration. Active vision systems do not rely on object visual features to obtain information, but take advantage of additional information, such as artificial controlled lights or pattern projection, to enhance the characteristics of the image scene.

The basic elements composing an active system, in fact, are a light source, a detection unit and a data processing unit. The main advantage provided by the artificial and controlled light source is the automation of object point identification, which is performed by the sensor itself. Secondly, active techniques are capable to acquire thousands (up to million) points in few seconds with a high resolution. Nowadays, several active sensors are available based on different measurement principles (Stjepandić et al., 2015, Guidi et al. 2010): range cameras based on laser beam and structured-light pattern projection, or time of flight systems (ToF) are common tools for scientific community and find many applications in several different fields. In the first case, the basic principle is the classical geometric triangulation that still provides good results in terms of accuracy. ToF systems instead rely on the measurement of the time the light signal sent on the scene goes back-and-forth. This is the principle of Light Detection And Ranging (LiDAR) and is based on the optical measurement of the sensor-to-target distance knowing the angles of the range measurement device.

Besides these well-known acquisition methodologies, it is important to mention among the active techniques the interferometry, deflectometry, and the active stereoscopy. Interferometry (specifically, white-light interferometry) is the most widely used method in industrial quality control. It is based on the phase shift analysis of two white light beams generating an interference pattern, and is suitable for smooth plane surface only. Deflectometry can be compared to structured-light scanning for 3D surface measurement of free-form shape. The working principle consists in projecting a geometrical pattern and checking for deviations in the reflected one; it is used to characterized optical surface (mirror and lenses) or for industrial application, such as car body parts acquisition. To conclude, active stereoscopy is similar to the passive one but takes advantage of a non-codified pattern: the aim is to facilitate the triangulation process through the help of the artificial features provided by the projected pattern.

Despite the recent development of the image-based techniques, active sensors are often preferred. Range-based data in fact are characterized by high reliability, and the capability to deliver detailed metrically-defined 3D models in short time overcomes the progresses reached by automated image-matching algorithms of image-based modelling. Next chapter will focus on triangulation scanners, structured-light technology and ToF systems.

## 2. IMAGE-BASED METHODS: CLOSE-RANGE PHOTOGRAMMETRY

### 2.1 Generalities: photogrammetric process

Photogrammetry is defined as a technique that takes advantage of methods of image processing in order to derive an accurate measurement of the shape of a given scene from one or more images of that scene taken from different points of view (Luhmann et al., 2019). The most common purpose of the photogrammetry is the reconstruction of the geometric characteristics of an object, such as dimensions, coordinates and location, in a 3D (three-dimensional) digital format. The photogrammetric process that leads to the reconstruction of the model encompasses the steps of (I) image recording/acquisition, (II) image interpreting/measuring and (III) image matching and 3D object reconstruction.

- (I) **Image recording/acquisition** is accomplished by a vision system governed by physical models that mimic the HVS. The vision systems used in photogrammetry have evolved over time thanks to the advancement of available technologies. Nowadays, we dispose of a wide variety of systems: *industrial cameras* (webcams, surveillance, industrial inspection, driver assistance cameras etc.), *high-speed cameras* capable to record fast-changing scenes, *scanning cameras* for high resolution results, *panoramic cameras*, *thermal imaging cameras* and *3D cameras*, that directly create a 3D point cloud. The use of these types of systems varies with the field of application: the vision systems discussed in this chapter are *digital cameras*, which are the most used devices for close-range photogrammetric surveys.
- (II) **The photogrammetric image measurement** has received a positive impulse since the image processing methods have developed and became available. This step can be also considered as a photogrammetric image interpretation: it involves mid-level image processing methods to derive spatial information from digital images for every image point. Among the image processing techniques used for image measurements, we mention the *edges extraction*, which is a peculiarity of the HVS for the object recognition; and the *segmentation of features*, particularly useful in the phase of automatic target recognition (section 2.4)
- (III) **Image matching** is the digital image processing method that consists in identifying object features on at least two images of the object and matching them. The features can be points, patterns or edges. This step of the photogrammetric process leads to the **3D surface reconstruction**. In image processing, there exists a wide range of matching techniques (Gonzalez and Woods, 2018). However, the strategies that are related to geometric surface reconstruction are essentially two, the structure-from-motion (SfM) relying on a feature-based matching, and the method based on the epipolar geometry and area-based matching. The application of one approach in respect to the other depends on the initial condition of the orientation parameters. (section 2.5)

This chapter focuses firstly on the acquisition of digital images through digital cameras and deepens the functioning system, and the calibration problem. Then, it shows the image processing methods used in digital close-range photogrammetry for the image measurement and the image matching, until the enunciation of the 3D object reconstruction methods.

## 2.2 Image acquisition: digital cameras

### Sensors

Digital cameras are the electronic imaging systems used for close-range photogrammetry. They acquire visual information through opto-electronic sensors: so as in the HVS the organ devoted to the acquisition of the scene are the eyes (see Chapter 1), in a digital camera the component responsible for the image acquisition is the sensor. The camera's eye directly provides an electronic image that can be either preprocessed in the camera, or transmitted to a host, processed and displayed on a monitor. The realization of opto-electronic sensors mimics the eyes behavior introducing some approximations in resolution and scanning temporally. Industrial sensors, in fact, have limited pixel resolution for both spatial dimensions and the image capturing is a discrete temporal scanning with defined frames per second.

Commonly, digital cameras use *CCD* (Charged Coupled Device) and *CMOS* (Complementary Metal Oxide Semiconductor).

- **CCD sensor:** is a solid-state imaging sensor mainly used in digital photogrammetric systems. Invented by American Bell laboratories in the 1960s, a CCD sensor is composed by silicon detector elements that are light sensitive. Detectors generate an electric charge proportional to the amount of incident electromagnetic radiation: when the incident illumination reaches the detectors, it is absorbed in a semi-conducting layer that is responsible for the creation of the charged particles (Figure 2.1). This photo effect uses photons to generate electrons, which are attracted by a positive electrode. Charges are then converted in current voltage and transformed by an analog to digital converter (A/D) in digital number, bit of information constituting the image (Luhmann et. al, 2019).

Detector elements are arranged in lines (1D sensors) or two-dimensional arrays (2D sensors) and the semi-conductors are coupled so as they can transfer the charged particles to the adjacent one. Depending on the layout of the 1D CCD sensors, we can distinguish between simple CCD line sensor and bilinear CCD line sensor. In the first case, each element of the sensor is directly connected to a read-out register, whilst in the bilinear CCD line one, the elements are resampled and connected to two read-out registers following an alternating behavior. A CCD line sensor can reach a length of 100 mm, since it can be composed by more than 12000 sensor elements with a spacing of ca. 4  $\mu\text{m}$  to 20  $\mu\text{m}$ . However, most medium and low-cost digital cameras use CCD sensors arranged in a 2D matrix, whose construction is more complicated than a simple line sensor for the reason that read-out process must be accomplished in two dimensions. Among the three arrangement of a CCD matrix sensor, namely *frame transfer* (FT), *full-frame transfer* (FFT) and *interline transfer* (IL), the most used on digital cameras devoted to photogrammetric image acquisition is the IL sensor. They alternate light sensitive and opaque CCD arrays, which are devoted to shift the charges. They are called “vertical transfer registers”, are covered by an optical metallic shield and allow the

lateral transfer of the charges that successively are shifted along the columns to the read-out registers.

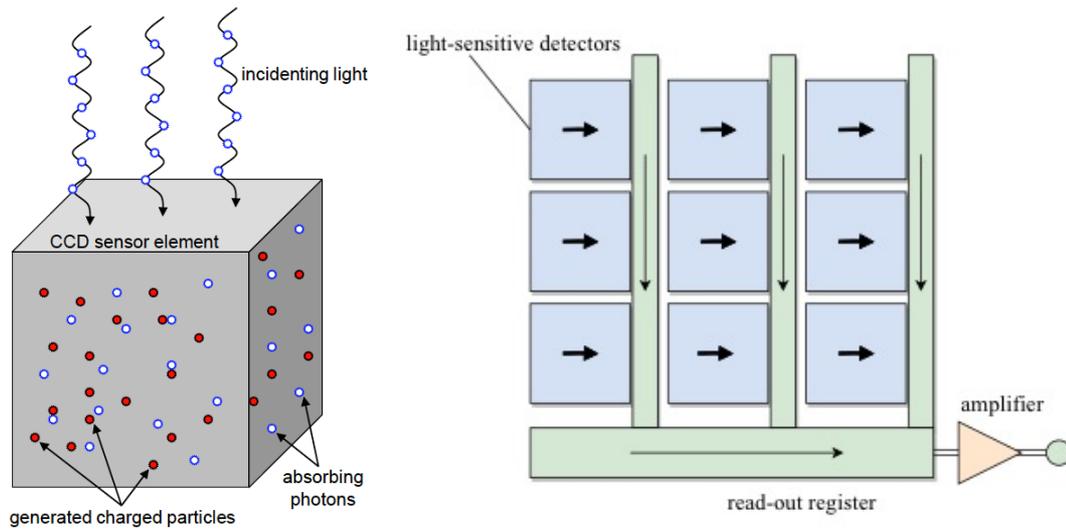


Figure 2.1. CCD sensor functioning: on the left, incident light (photons) is converted into charge. (Luhmann et al., 2019). On the right, read-out technique of a IL CCD sensor.

These types of sensors, have the advantage to allow both image acquisition and readout, which are performed in parallel facilitating faster frame rates. Thanks to the fact that the pixel area storage cannot be hit by light, IL sensors avoid the image smear but, for the same reason, are less light sensitive (only about 25% of the total area is covered by detection sensors). This implies that IL is not the best sensor technology to use for applications that lack of light. However, the reason of their employment in commercial cameras resides in the small size area, which results in a relatively inexpensive manufacturing (Guidi et al., 2010; Luhmann et al., 2019).

- **CMOS sensors:** *Complementary Metal Oxide Semi-conductor* technique, despite being not new, has developed significantly in recent times. CMOS sensors now are widely used for computer processors, memory chips and digital circuits. The basic principle is the photovoltaic effect (Hornberg, 2017): differently from CCD sensor that holds and transfer the charge, in CMOS sensors the conversion of the charge to voltage is realized inside the pixel and is possible thanks to an integrated amplifier and digitizer unit attached to each pixel element. CMOS sensors have higher level of integration and lower power consumption than CCD sensors. On the other hand, they are slow in *read in/out* operation and requires higher energy.

## 2.3 Camera calibration

Nowadays, the employment of metric or semi-metric cameras for close-range applications is barely disappearing, while commercial non-metric cameras are displacing them. Multiple reasons there exist: commercial cameras can reach higher image resolution, they are compact, portable, easy manageable and relatively cheap. They have facilities in storing and transferring images files and the image acquisition is easy and direct. The images are generated in digital format, without the employment of digitizing films or paper prints. Nevertheless, commercial amateur cameras do not have any calibration certificate, and for these reasons, accurate calibration and orientation procedures are necessary for the extraction of 3D metric information characterized by precision and reliability. The aim of camera calibration is to determine the geometric camera model, i.e. to know the internal orientation (IO) parameters: the principal distance, the image coordinates of principal point, and lens distortion parameters.

Over the years, heterogeneous camera calibration algorithms have been developed, which typically rely on perspective or projective camera models. Moreover, according to the reference object type, time and location of the process, three approaches can be distinguished: laboratory, test-field and self-calibration. Since the *laboratory calibration* refers to old metric cameras, here we refer to the *test-field calibration* and the *self-calibration*. These two approaches are the most popular calibration methods used for non-metric cameras. *Self-calibration* is nowadays a basic and regularly employed approach within photogrammetry process: it was introduced for the first time in the early 1970s, and is currently applied in high-accuracy close-range photogrammetry within the more general *bundle adjustment* solution.

### The central projection and collinearity equations

The photogrammetric calculation is based on the *central projection* fundamental to reconstruct the location of a point in the space from the correspondent image coordinates. The *central projection* rules allow imaging an object on a plane, projecting it from a perspective center placed at a finite distance (Figure 2.2). Assuming a rigorous central projection, the result is that the object points  $\mathbf{X}_P$  and the corresponding image points  $\mathbf{x}'$  are located on the same line passing through the perspective centre  $\mathbf{X}_0$ . Object points and image points can therefore be obtained from the position vector to  $\mathbf{X}_0$  and from  $\mathbf{X}_0$  to the object point. The knowledge of the position of the perspective center  $\mathbf{X}_0$  is the *internal orientation* (IO) and establishes the relationship between the image plane and the projection center of the camera. Its reconstruction is defined by the image coordinates of the principal point  $H'$  (measured in an image-based coordinate system) and the principle distance  $c$ .

Considering an ideal pinhole camera, all the image rays pass through the perspective centre O. The IO most important parameter is the principal distance  $c$  that defines the distance between the image plane and the perspective centre. This optical model, where the object points, the image point and the projection center are aligned, satisfies the *collinearity condition*. However, a real digital camera differs from the pinhole camera model and the interior orientation must include additional parameters that take into account lenses, the instability of camera housing and the optical axis variations. Camera interior orientation including these deviations from collinearity condition, called distortions, is determined by calibration procedure.

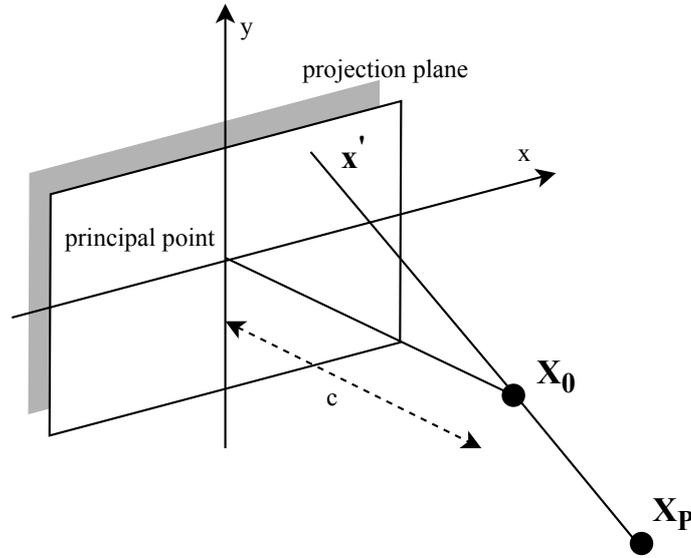


Figure 2.2. Geometry of central projection.

To know the spatial geometry of an object is also necessary to know the position and orientation of the camera with respect to a global reference system. This information constitutes the *external orientation* (EO) of the image. The external parameters specify the coordinates of the perspective centre and three angles expressing the image coordinate system rotation with respect to the global system. These six parameters are determined indirectly, after the measurement of object point image coordinates.

The transformation of the object coordinates into the corresponding image coordinates, is described by the *collinearity equations*. They derive from the spatial similarity transformation and specifically define the projective transformation between the image space and the object space as function of interior orientation parameters and exterior orientation parameters on one image.

$$\mathbf{X}_P = \mathbf{X}_0 + m \cdot \mathbf{R} \cdot \mathbf{x}'$$

$$\begin{bmatrix} X_P \\ Y_P \\ Z_P \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + m \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \quad [2]$$

Where:

$\mathbf{X}_P$  = object coordinates

$\mathbf{X}_0$  = perspective centre coordinates

$m$  = scaling factor

$\mathbf{R}$  = rotation matrix

$\mathbf{x}'$  = image coordinates

Equation [2] describes the projection of an image point into a corresponding object point: here the scale factor  $m$  represents the relationship between the object distance  $h$  and the principal distance  $c$ .

$$m = \frac{h}{c}$$

Since the image coordinates are the system unknown, the [2] system is inverted, and principal point  $H'(x'_0, y'_0)$  and image distortion parameters (through the correction term  $\Delta\mathbf{x}'$ ) are added. The collinearity equations follows:

$$x' = x'_0 + z' \cdot \frac{r_{11} \cdot (X - X_0) + r_{12} \cdot (Y - Y_0) + r_{13} \cdot (Z - Z_0)}{r_{13} \cdot (X - X_0) + r_{23} \cdot (Y - Y_0) + r_{33} \cdot (Z - Z_0)} + \Delta x' \quad [2.2]$$

$$y' = y'_0 + z' \cdot \frac{r_{21} \cdot (X - X_0) + r_{22} \cdot (Y - Y_0) + r_{23} \cdot (Z - Z_0)}{r_{13} \cdot (X - X_0) + r_{23} \cdot (Y - Y_0) + r_{33} \cdot (Z - Z_0)} + \Delta y'$$

The object points  $(X_P, Y_P, Z_P)$  are transformed into unique image points  $(x', y')$  as function of  $x'_0, y'_0, c, \Delta x', \Delta y'$ , which are the interior orientation parameters, and  $X_0, Y_0, \omega, \varphi, k$  (rotation angles, elements of the rotation matrix  $\mathbf{R}$ ), namely the exterior orientation parameters of one image (Luhmann et al., 2019). The collinearity equations, in a more general sense, describe image formation inside a camera following the central projection geometric principle. They are the fundamental of the photogrammetry and are used for example in the spatial intersection, space resection, orthophotos generation and bundle triangulation.

### Bundle adjustment and self-calibration

In high accuracy close-range photogrammetry, digital cameras are calibrated using the analytical model of the collinearity equations within the *bundle adjustment* solution. This indirect method derives interior parameters of a chosen camera model from image coordinate observations. It uses the modified version of the collinearity equations to take into account systematic errors coming from non-modeled effects of the image acquisition process. As described above, the errors are usually represented as  $\Delta\mathbf{x}'$  and  $\Delta\mathbf{y}'$  and are introduced as additional parameters (APs) that correct image coordinates. Actually, in photogrammetry, there exist various sets of APs and the most used is the one proposed by Brown (Brown, 1971). It includes 10 interior orientation camera parameters ( $\Delta x$ ,  $\Delta y$ ,  $\Delta c$ ), the pixel shape uncertainty (shape factor  $S_x$ ), the factor modeling the non-orthogonality of the image coordinate system (shear factor  $A$ ), the symmetrical radial lens distortion ( $k_1, k_2, k_3$ ), and the decentering lens distortion ( $p_1, p_2$ ). This is the most employed method, since it is considered the most flexible and accurate one (Brown, 1976; Granshaw, 1980; Triggs et al., 2000) and is also known as *self-calibration*.

*Bundle adjustment* method is basically a global minimization of the reprojection errors: since the system of equations is non-linear, it is necessary to introduce approximate values for all the unknown parameters. Their computation is made iteratively by the least square method, so that the square deviations are minimized at the image coordinates observed. Moreover, the quality of the imaging system is given thanks to the computation of the standard deviation of all parameters (Hornberg, 2017). The *bundle adjustment* approach thus allows the simultaneous calculation of the unknown object coordinates, exterior orientation, and interior orientation parameters of the imaging system. The data flow for a *bundle adjustment* process is shown in Figure 2.3. The image coordinates, which are calculated by manual or automated image measuring system, are corrected by the *bundle adjustment* calculation that at the same time performs a statistic error analysis on the initial values of

the unknowns. The final output of the bundle adjustment are the exterior and interior orientation parameters.

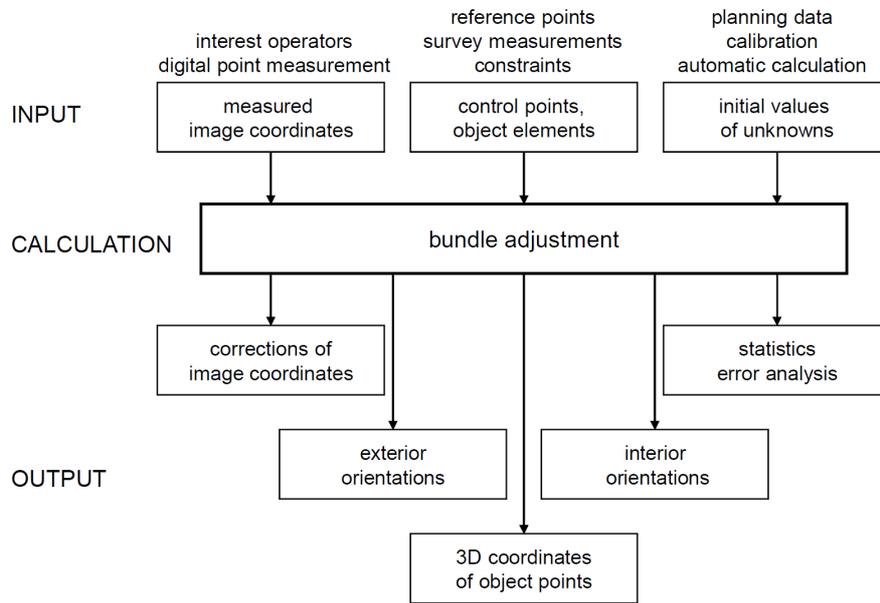


Figure 2.3. *Bundle adjustment* data flow.

Here we consider the application of *bundle adjustment* method in a multi-image processing where an unlimited number of spatially distributed images is considered as the starting point of the process. The measured image points and the projection centre of each image define a spatial *bundle of rays*. As Figure 2.4 shows, every object points lies on the direction vector corresponding to every measured image points. The vector describes the spatial direction from the projection centre to object points that can be located along the image ray. The object points can be absolutely determined only knowing additional geometric constraints, like a spatial direction or an object plane which, intersecting the ray, define the correct position in the space hence defining a dense network. The *bundle triangulation* method then, will perform the orientation of images and the calculation of the 3D object point location in space.

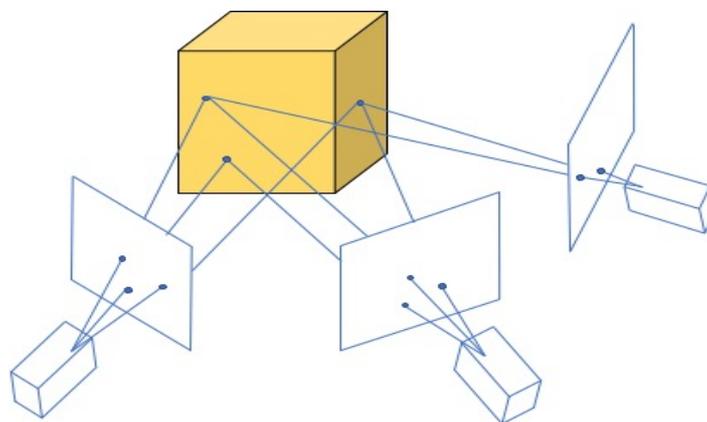


Figure 2.4. *Bundle adjustment* schematization.

To conclude, solving a *bundle adjustment* with self-calibration means to estimate the APs in the extended collinearity equations and, at the same time, the interior and exterior orientation parameters for every image of the set. The object spatial coordinates of the points will be estimated through the collimation of homologous points in each image. The equations derived from the combination of the equations written for every point in all images is a redundant system that have its solution.

### Direct Linear Transformation (DLT)

The process of calculating the exterior orientation parameters can be led back to the problem of single images orientation. Since direct orientation methods are often impossible to employ, it is common to make use of indirect methods for the calculation of exterior orientation. The Direct Linear Transformation (DLT) is the most popular method among the ones based on projective relations: it provides a direct solution requiring six reference points without the need of approximate values. Analytically, DLT model is a function capable to set a biunivocal relation between image space and object space, regardless of the optical system acquisition technology. The mathematical relation is a linear type relation in the unknown  $L_i$ , and for this reason it does not require any approximate value. The equations describing the DLT model are based on the collinearity equations, modified with an affine transformation of the image coordinates. An image coordinate system is therefore not required.

The DLT equations follows:

$$x = \frac{L_1X + L_2Y + L_3Z + L_4}{L_9X + L_{10}Y + L_{11}Z + 1} \quad [2.3]$$

$$y = \frac{L_5X + L_6Y + L_7Z + L_8}{L_9X + L_{10}Y + L_{11}Z + 1}$$

In [2.3]  $x$  and  $y$  are image coordinates,  $X, Y, Z$  the 3D coordinates of the reference points, and  $L_1, \dots, L_{11}$  are the DLT independent parameters that need to be estimated (Luhmann et al., 2019). They are related to the acquisition of the photogrammetric data and are function of interior and exterior parameters: for this reason, the equations [2.3] can be formally compared to the collinearity equation, even though they are less general and can be used in few application fields. The DLT is described here in fact, because it is mostly used for the close-range photogrammetry. In general, the employment of DLT has several benefits. They can be applied to images with arbitrary coordinate system and the elaboration is fast because the linearity of the equations allows skipping the estimation of unknown initial values. However, there are also some drawbacks. DLT has an excess of parameters if interior orientation parameters are known, and, if the reference points are located on the same plane, the system of equations become weakly conditioned or singular. DLT cannot detect image coordinates or reference points coordinates errors, providing possible false parameters and, to conclude, the minimum number of reference points cannot always be ensured in real applications.

## 2.4 Image measurement

### 2.4.1 Edges extraction

Edges extraction is a basic tool in image processing and CV, and is particularly useful for the feature detection and extraction (Ziou and Tabbone, 1998; Hornberg, 2017; Luhmann et. Al, 2019). The primary aim of edge extraction is to find a set of connected curves that identify the boundaries of objects within images. The edges extraction algorithms work detecting discontinuities in brightness. The aim is to capture important changes in image structure: considering a bi-dimensional image, the edges extraction operation is based on the fundamental structure of the gray-value edges. The changes to be detected are the changes in gray values along the physical object edges. Assuming that the image contains no noise, that is to say that the gray values are constant in the object and in the background, the edges can be defined as areas in which gray values (grey-value edges) change significantly. The definition of edge can be formalized as follow: regarding the image as a 1D function  $f(x)$ , the grey value changes if  $|f'(x)| \gg 0$  and  $|f(x)|$  is locally maximal (non-maximum suppression), that is to say that in this point the second derivative vanishes ( $f''(x) = 0$ ). It follows that in 1D condition, the inflection points of  $f(x)$  corresponds to edges. There exist several filters facilitating the edge detection.

One of the most known is the Roberts Cross operator, a first order differential filter that approximates the discrete first derivative of the image function  $f(x)$  and thus performs simple 2D spatial gradient measurement on the image. Another gradient filter is the Sobel operator, which consists of a pair of 3x3 convolution kernels that approximates the first derivative in  $x$  and  $y$  (Sobel, 2014). This operator is slower than the Roberts Cross one, but is less sensitive to noise. Figure 2.5 shows the edges extraction with Roberts and Sobel operator separately.

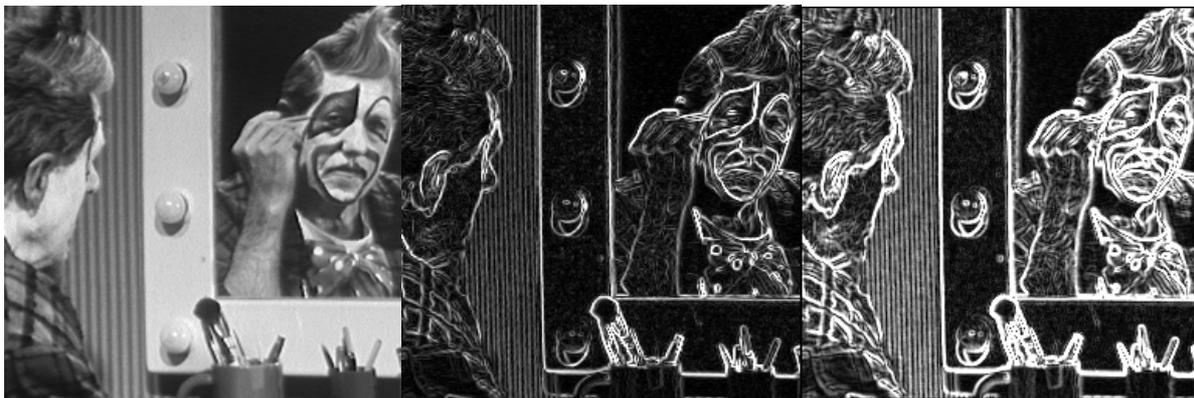


Figure 2.5. From left to right: original image, the image after the Roberts Cross filter application and after Sobel algorithm. Sobel operator implies an increased smoothing: therefore, the edges extracted by Sobel are thicker compared with the Robert Cross output.

### 2.4.2 Image segmentation

The image segmentation procedure is a mid-level image processing method that aims to extract regions from an image. The regions correspond to features of particular interest in the image, namely the objects to be recognized and measured. Formally, the segmentation procedure involves images as input and returns regions or contours as output. In photogrammetric applications, the output regions to be searched and measured are often the targets (when no information about their position is known): for this purpose, different segmentation methods are available, depending on image scales, distortion, and lighting and on target characteristics (retro-reflective type, printed etc.).

The simplest segmentation algorithm involves the image thresholding procedure. It is an adaptive binarization of image that selects all points in a region of interest (ROI) that are included within a defined range of grey values. It is used in general to separate objects and background and is particularly suitable in case of retro-reflective and LED targets. The adaptive binarization, although being simple and fast, provides segmented regions which, in turn, contain multiple objects. These sub-regions need to be segmented individually, and this is possible by the extraction of connected components, namely connected set of pixels. Among several types of pixel connectivity, image segmentation takes advantage of the 4- and 8-connectivity (Figure 2.6).

- The 4-connectivity: each pixel has four connected pixels, those that have an edge in common.
- The 8-connectivity: each pixel has eight connected pixels, including also the diagonally adjacent ones.

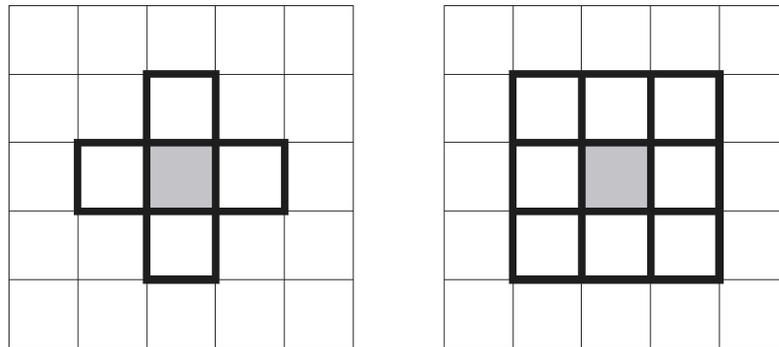


Figure 2.6. 4-connectivity (left): starting from a selected pixel (the grey pixel) the 4-connectivity individuates the four pixels that share with the initial one at least an edge. 8-connectivity (right): considers also the pixels that are diagonally adjacent to the initial one.

The connectivity analysis uses a sequential process that examines the left and the upper three neighboring pixels at the current position. If the grey value threshold is exceeded in one of these neighbors, the pixel is classified as part of the region already detected. Otherwise, the process comes up with a new region. In order to avoid inconsistencies, it is a good practice using opposite connectivities on the foreground and background (Hornberg, 2017). This analysis procedure can also be performed on binary represented regions and it will return as output a label image (labeling). The regions resulting from the connectivity analysis can be analyzed with respect to their shape and size. A common criterion is based on connected pixel number. However, in presence of circular targets, the standard deviation of a best-fit ellipse from the contour or the ratio of semi-axes is suitable. An interesting application of this analysis is to remove unexpected resulting regions caused by noisy segmentation by setting a threshold on the area. After the operation mentioned above, the resulting regions are indexed and stored.

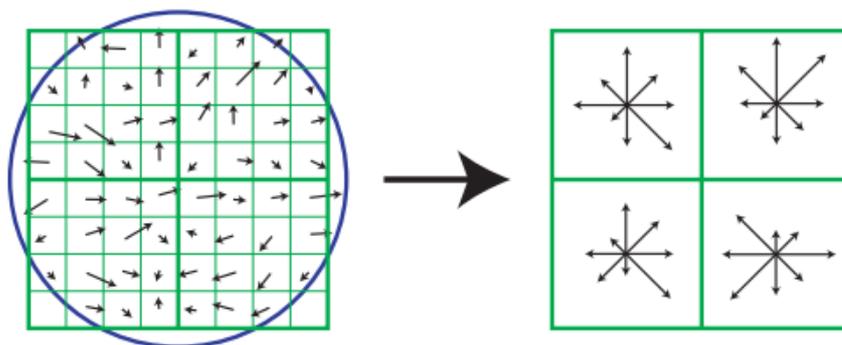
The image segmentation in photogrammetric application is a central issue, and the fully automated segmentation in particular can lead to successful results in the target recognition phase.

### 2.4.3 Feature extraction

Feature extraction consists of a series of operation governed by algorithms that aim to determine one (or more) characteristic quantities from the regions of an image. In fact, even though edges and regions are very useful, for example, during the target recognition phase, they only lead to raw description of the segmented entities. Photogrammetry takes advantage of feature extraction method to search for image points to be used during the image-to-image matching: the operators involved in this process are called *interest operators* and allow to perform the detection of distinctive image points. The operators search for interesting features that satisfy the requirement of individuality, invariance (radiometric and geometric), robustness, rarity and applicability.

In CV there exist several criteria for determining the presence of distinctive features, which are realized by means of mathematical operators. The algorithms can analyze, within a local window of the image, the local variance, auto-correlation function, self-similarity, grey-value surface curvature, gradient sums or local grey value making a comparison. To each of these criteria, correspond an operator. The most used are Forstner, SUSAN, FAST, SIFT and SURF (Forstner and Wrobel, 2016; Wenfei et al., 2017; Luhmann et al., 2019).

The most used algorithm in photogrammetric field is *SIFT operator* (Scale Invariant Feature Transform), prosed by David F. Lowe in 1999. It governs a procedure for the feature detection structured in four steps: after transforming the input image in image pyramids and smoothing each level by a Gaussian filter, it starts extracting edges and their extrema. The procedure searches for local extremum within 8-neighborhood of the current image and 3x3 neighbors of the upper and lower pyramid. The individuated features are taken as initial key points for the second step, which eliminates thin features with low contrast or smallest edges based on gradient magnitude and surface curvature. The gradient is the key characteristic for the extrapolation of the orientation angle of the feature: the third steps consists in storing the weighted values of gradient magnitude and direction in a 10 classes histogram, from which individuating the main orientation angle of the feature. A refinement of this measurement is then performed by calculating the gradient magnitude in smaller 4x4 pixels blocks and weighting it based on the reciprocal of the distance from the feature point. By calculating the gradient histogram in 8 direction, seed points are obtained, i.e. a descriptor is set. The result is a normalized SIFT feature vector (Figure 2.7).



2.7. SIFT feature descriptor (Lowe, 1999).

## 2.5 Image matching and 3D object reconstruction

In digital photogrammetry, the image matching is the fundamental process that leads to the 3D surface reconstruction of the scene. The methodology consists in individuating corresponding entities in two or more images, the object features such as points or edges, and performing a matching between them. However, in CV the automatic matching of corresponding image features has been always an issue and, even in recent times, is one of the most researched topic. The first problem to face is the analysis of the scene, which in the majority of the cases, is not structured. The unstructured scenes, devoid of any arranged objects (e.g. targets), often hinder the correspondence analysis presenting several problems of occlusion, ambiguity, instability and noise sensitivity. Thus, the understanding of the spatial scenes is an ill-posed problem, that can be solved assuming some preconditions related to the object and image acquisition. Over the years several algorithms have been developed. For its structure, each algorithm is capable to manage and work on different type of data (Jayanthi and Sreedevi, 2018); the most successful and used matching methodologies for geometric surface reconstruction are the *feature-based* and *area-based* matching: they respectively satisfy two common initial orientation conditions of a set of images:

- Known interior orientation, unknown exterior orientation parameters (feature-based);
- Known interior and exterior orientation parameters (area-based).

In the next paragraphs both matching approaches are analyzed describing differences between them.

### 2.5.1 Feature-based matching and Structure from Motion approach (SfM)

The method based on the feature matching relies on the independent identification and extraction of distinctive elements from multiple images that are not yet oriented, namely it is based on the feature-extraction operation (2.4.3). During the feature extraction preliminary step, punctual, linear and areal features are extracted by the feature detectors and edge operators, whilst target are detected taking advantage of segmentation methods (2.4.2). The extracted features are grouped in attributing structures where they are characterized by:

- Geometrical parameters (orientation, curvature, center);
- Topological relationships;
- Radiometric characteristics;
- Texture;
- Point numbers.

A preliminary list contains all the assumed corresponding features common to all images. The next step is the proper *feature-based matching*, that elaborates the final couples list containing the calculated refined corresponding features. The algorithm firstly attempts to identify as many corresponding features as possible in all images through similarity functions; then finds parameters of relative orientation of a selected suitable pair of images through a robust calculation (RANSAC methods), and finally connects all other images by corresponding points. The exterior orientation parameters (EO) are calculated for all images after the camera calibration step through the bundle adjustment approach (2.2.2). The estimation of the 3D coordinates of the feature points and the generation of a sparse point cloud follows. This procedure is commonly and widely known as *Structure from Motion (SfM)*.

***Structure from motion (SfM)***

Born in 1979 (Ullman, 1979), the SfM approach enables the direct estimation of both internal and external parameters (IO and EO) and the automatic reconstruction of the three-dimensional object or structure from the motion of a single camera around the object. Typically, the workflow uses sequentially: a combination of image processing algorithms aiming to the feature detection and matching, robust orientation RANSAC-based methods, bundle adjustment with self-calibration, and a multi-view stereo analysis for the dense model reconstruction (Figure 2.8).

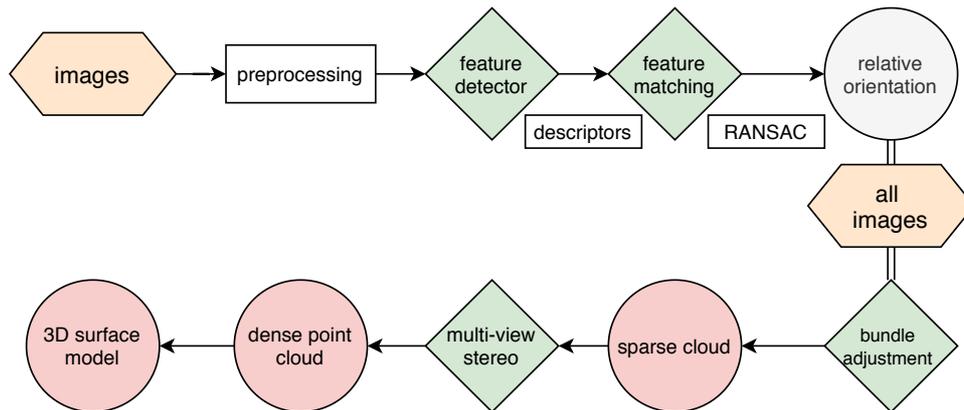


Figure 2.8. SfM workflow, from a set of images to the 3D surface model.

After the feature detection, which is often SIFT-detector based (2.4.3), and the feature-based matching, a RANSAC-based relative orientation step is conducted. RANSAC (random sample consensus) algorithm permits the iterative estimation of the orientation parameters through the repeated calculation of a function using the minimum number of observations. The tolerance parameter  $d$  affect considerably the success of the algorithm, that search for valid observations which must be consistent with the calculation.

Once the best solutions are identified, they constitute a consensus group, while the outlier observations are rejected. The relative orientation based on RANSAC algorithm starts from an initial pair of images, which constitutes the starting model for the subsequent orientation of all remaining images. At this point of the process, the bundle adjustment processes the estimated object coordinates that compose the sparse point cloud, a first approximate distribution of 3D surface points. From the 3D point cloud a digital surface model (DSM) can be created through the Multi-View Stereo (MVS). MVS is the matching extended to any number of images: it identifies an image (reference image) and searches for overlapping images to matched. If  $n$  is the number of images,  $n$  disparity maps are calculated and then normalized.

### 2.5.2 Area-based matching

The method, sometimes called *signal-based method* or *template matching*, relies on the similarity computation between a patch on one image and the corresponding patch on the second coupled image. The patch is known as *template*. Specifically, the comparison involves the grey values  $f(i, j)$ . As Figure 2.9 shows, the procedure consists in fixing a *template* in one image and analyzing a *search window* in the second one, where the grey values  $g(i, j)$  are compared in order to find the corresponding patch.

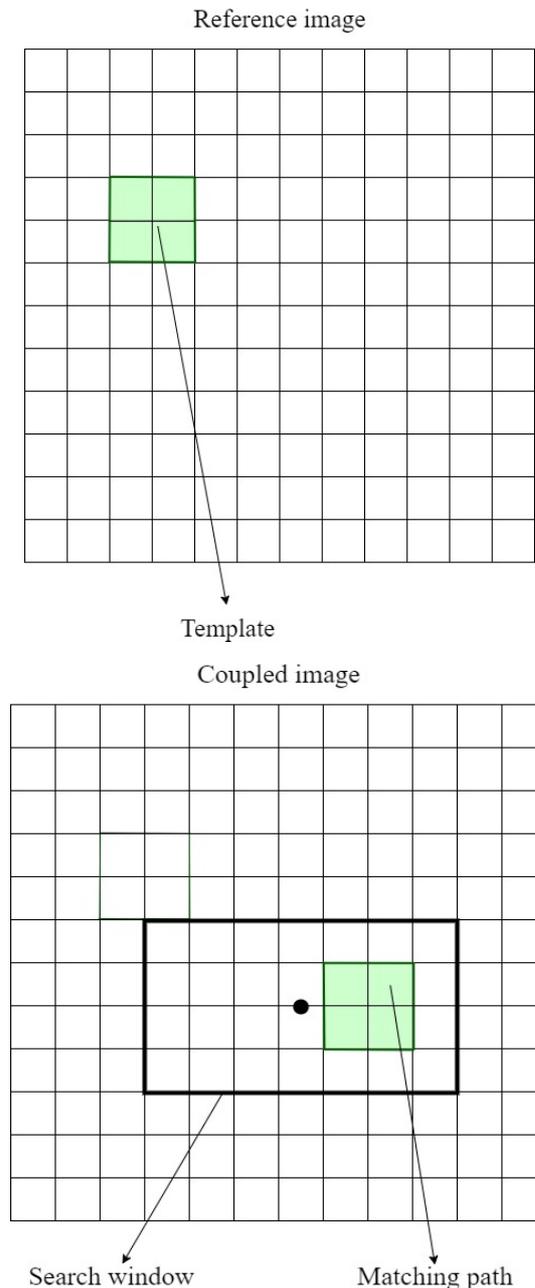


Figure 2.9. Area-based matching approach: in green the patch in the reference and coupled image. The search window contains the matching patch.

One of the simplest criterion for the identification of the best correspondence between *template* and *matching window* is based on the correspondences analysis relying on the measurement of the similarity through the value of the correlation coefficient  $\rho$ . The coefficient is defined as:

$$\rho = \frac{\sigma LR}{\sigma L \sigma R}$$

Where:

$\sigma L$  is the standard deviation of the density function  $g_L(x, y)$  related to the image template;

$\sigma R$  is the standard deviation of the density function  $g_R(x, y)$  related to the matching window;

$\sigma LR$  is the covariance of the two densities;

The correlation coefficient is calculated for each line and row of the matching window; than the algorithm searches for the position inside the window that maximizes it.

The most used method to search the best matching window is the *least squares matching* (LSM). Even though less simple than the correlation method described above, it is more versatile and efficient (Grün, 1985). It considers, as in the previous case, the grey values of the template  $f(i, j)$ : the similarity is calculated through the geometric and radiometric transformation between a reference image and coupled image such that the least-squares sum of grey values differences between both images is minimized [2.4]. The images to be compared are usually a stereo pair: the reference image can be a window in a real image or can be synthetically generated to be used as a template; in this case, the image matching method is called *least squares template matching*.

$$\sum_i \sum_j (f(i, j) - g(i, j))^2 = \min \quad [2.4]$$

The mathematical model follows:

$$f(x, y) - e(x, y) = g(x, y) \quad [2.5]$$

Where:

$e(x, y)$  is a noise component

Due to the linear transformation expressing the radiometric and geometric transformation of grey values at position  $(x, y)$ , the reference image  $f_i$  can be expressed as follow:

$$f_i(x, y) - e_i(x, y) = r_0 + r_1 \cdot g_i(x', y') \quad [2.6]$$

$$x' = a_0 + a_1x + a_2y$$

$$y' = b_0 + b_1x + b_2y$$

$$i = 1, \dots, n$$

$n$  : window size

$a_0, b_0$  : translation parameters of the relative shift between images

From the linearization of the equation, we obtain:

$$f_i(x, y) - e_i(x, y) = g^o(x, y) + g_x da_0 + g_x x da_1 + g_x y da_2 + g_y db_0 + g_y x db_1 + g_y y db_2 + dr_0 + dr_1 g^o(x, y) \quad [2.7]$$

$$g_x = \frac{\partial g^o(x, y)}{\partial x}$$

$$g_y = \frac{\partial g^o(x, y)}{\partial y}$$

The LSM model is generally valid for standard cases, but can also be extended or simplified depending on the similarity between images and the accuracy of initial values. This approach, even though introduced for a couple of images patches, can be extended to multiple images resulting particularly useful for simultaneous matching of multiple points in multiple images. Additionally, geometric constraints can be set up between images: in this case, the correspondence analysis is based on epipolar geometry.

### Multi-image least-squares matching: MGCM approach

The LSM approach can be extended by the simultaneous matching of one point in multiple images. Considering a reference image  $f(x, y)$  and  $m$  searching images  $g_i(x, y)$ , the equation [2.5] can be extended as follow:

$$f(x, y) - e_i(x, y) = g_i(x, y) \quad [2.8]$$

$$i = 1, \dots, m$$

This relation enables the simultaneous determination of all matches of a point searching on multiple images, but does not imply that the all the homologous points belong to one common object point. The addition of this constraint must suggest the uses of collinearity equations: the condition is that, knowing the orientation parameters, the homologous image rays must intersect optimally at the corresponding object point. One effective method that combine the intensity observation equations with the collinearity condition and, for the following reasons, is very often used in close-range photogrammetry, is the Multi-photo Geometrically Constrained Matching (MGCM). Introduced in 1985 by Grün, besides introducing multiple images, MGCM overcomes the LSM approach limits of matching reliability and dimension of search window by using a geometric constraint. It reduces the multiple solutions given by a redundant texture and directly computes 3D object points coordinates and their accuracies (Previtali et al., 2013). The mathematical model of MGCM approach is described by observation equations expressing the grey level matching [2.8] and extended collinearity equations [2.2], which are used as geometrical constraints for the determination of object point coordinates.

Considering an image point P in image k, the coordinates of P are given by:

$$x'_{Pk} = \frac{1}{m_{Pk}} \cdot \mathbf{R}_k^{-1} \cdot (\mathbf{X}_P - \mathbf{X}_{0k}) \quad [2.9]$$

$$y'_{Pk} = \frac{1}{m_{Pk}} \cdot \mathbf{R}_k^{-1} \cdot (\mathbf{Y}_P - \mathbf{Y}_{0k})$$

where:

$m$  is the scaling factor;

$R_k$  is the rotation matrix;

$\mathbf{X}_{0k}, \mathbf{Y}_{0k}$  are the coordinates of the vector expressing the perspective centre coordinates.

Based on initial values  $(x'_p, y'_p)^0$ , the previous equations can be written as follow:

$$\Delta x'_{Pk} + F_{Pk}^X + x'_{Pk}{}^0 = 0 \quad \Delta y'_{Pk} + F_{Pk}^Y + y'_{Pk}{}^0 = 0 \quad [2.10]$$

$\Delta x'_p$  and  $\Delta y'_p$  are the shift coefficients (unknown).

Assuming that EO and IO of all images are known, the equations can be linearized at initial values of unknowns  $X, Y, Z$ :

$$\Delta x'_{Pk} + \frac{\delta F_{Pk}^X}{\delta X_P} dX_P + \frac{\delta F_{Pk}^X}{\delta Y_P} dY_P + \frac{\delta F_{Pk}^X}{\delta Z_P} dZ_P + F_{Pk}^X{}^0 + x'_{Pk}{}^0 = 0 \quad [2.11]$$

$$\Delta y'_{Pk} + \frac{\delta F_{Pk}^Y}{\delta X_P} dX_P + \frac{\delta F_{Pk}^Y}{\delta Y_P} dY_P + \frac{\delta F_{Pk}^Y}{\delta Z_P} dZ_P + F_{Pk}^Y{}^0 + y'_{Pk}{}^0 = 0$$

MGCM, combining the intensity observation equations [2.8] with the collinearity equations, allows the simultaneous determination of image and object coordinates (Baltsavias, 1991). The approach consists in the combination of LSM and original grey-values observations, but can also be integrated with additional observation equations, from the simultaneous calculation of object coordinates to the setting of other geometrical constraints ( $X, Y = \text{constant}$  or  $Z = \text{constant}$ ). Besides the cited common restrictions, several other additional constraints can be introduced to accomplish the task of edges extraction, object contours determination, and point grid measurement. In each case, the additional observations are weighted, in order to give them the proper effect on the model.

### Image matching based on epipolar geometry

Epipolar geometry describes the geometric configuration of a stereo vision referred to an object point  $P$ . Considering Figure 2.10, the geometric elements of the stereo configuration are:

- $b$ , the base;
- $r$  and  $r'$ , the projected rays from each perspective centre to  $P$ ;
- the **epipolar plane** (or **basal plane**), the portion of space defined by  $r$ ,  $r'$  and  $b$ ;
- $k'$  and  $k''$ , the **epipolar lines**, defined by the intersection of the epipolar plane and the image planes.

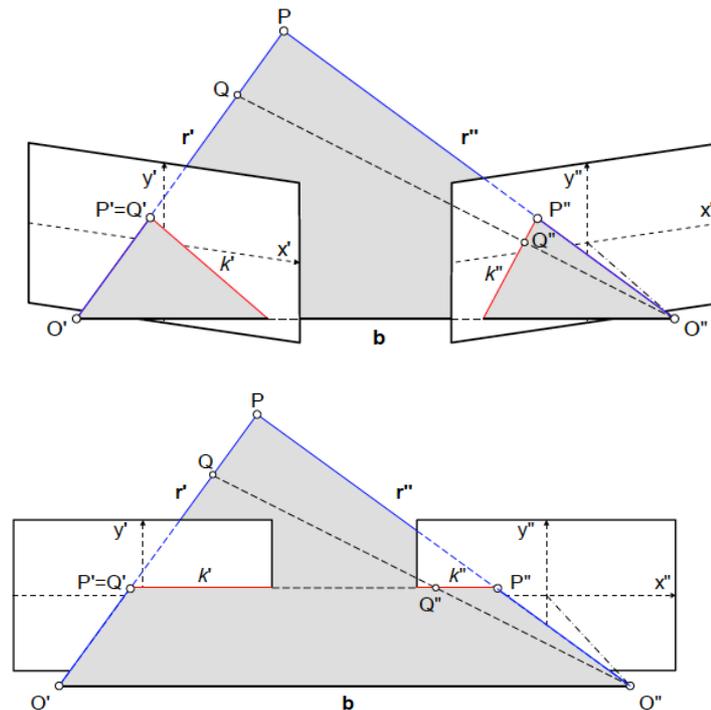


Figure 2.10. Epipolar geometry: in convergent images, the epipolar lines  $k'$  and  $k''$  are sloped at an angle to the  $x$ -axis (above); in the case of normal case, the epipolar lines are parallel to the  $x$ -axis (below). (Luhmann et al., 2019).

The use of epipolar geometry as image matching technique relies on the assumption that, if the relative orientation parameters of images are known, the position of epipolar lines can be easily calculated. In this way, the search space for corresponding point in the images is significantly simplified, since it can be reduced to an epipolar line. Hence, the application of image matching based on epipolar geometry is particularly effective in close-range photogrammetry application, where image orientation is performed separately, for example with SfM approach.

This method however has some weaknesses:

- Convergent images implicate higher computational effort than normalized stereo images: in the first case in fact, since the epipolar lines are sloped, the calculation of the image point  $P$  in both left and right images require considering possible distortions, which would curve the epipolar line.

- The uncertainty related to the orientation parameters could result in an extension of the search space from the epipolar line to a band which contains it. The band width  $\varepsilon$  increases with the uncertainty of the orientation and the measurement quality of  $P'$  (Figure 2.11)

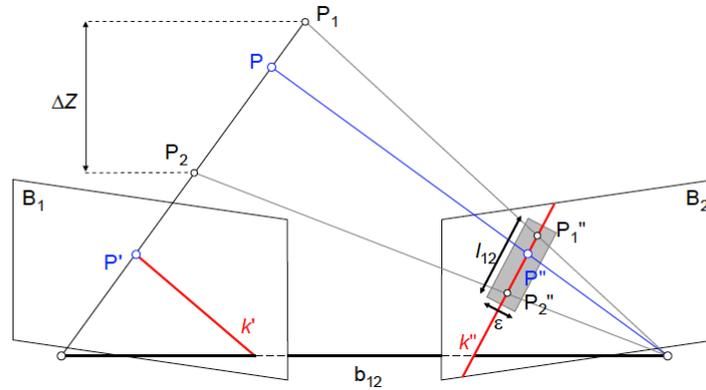


Figure 2.11. Image matching using epipolar lines. In grey, the area corresponding to the search space: its width is  $\varepsilon$  and depends on the orientation uncertainty and measurement quality of  $P'$ . The length  $l_{12}$  is function of the depth  $\Delta Z$  in object space (Luhmann et al., 2019).

- The total number of ambiguities  $N_a$  for each image pair is related to the number of image points  $n$ , the principal distance  $c$ , the width of the search space  $\varepsilon$ , the base length  $b_{12}$ , the object depth  $Z_{max} - Z_{min}$  and image area  $F$ .

$$N_a = (n^2 - n) \frac{2 \cdot c \cdot b_{12} \cdot (Z_{max} - Z_{min})}{F \cdot Z_{max} \cdot Z_{min}}$$

### 2.5.3 3D object reconstruction

The 3D point clouds derived from the image matching procedures described above represent the first product for the object reconstruction. Since the points do not have any structured organization in the space, they are known as *unstructured point clouds*. In general, a surface is modeled using a mathematical function in space domain which is very frequently of polynomial type: depending on the number of terms used, the surface model can assume a more and more complex shape. The *triangle-based surface modeling* approach relies on the inspection and determination of the first three terms (two first-order and zero-order terms), that form a planar surface, namely a spatial triangle which represent the constitutive element of the entire tilted planar surface. If the surface reconstruction involves also the fourth term, the resulting surface is referred to as a bilinear surface, represented as a simplified regular square grid (*grid-based surface modeling*). In both cases and in general, the process that start from the unstructured point clouds and leads to the surface reconstruction, is called meshing.

However, a distinction between two surface conditions must be set (Figure 2.12):

- Surfaces described by a function  $Z = f(X, Y)$  are *2,5D* (or  $2^{1/2}D$ ) *surfaces*. To each  $Z$  value, it is associated one and unique point of XY plane and the distance from the XY surface points is derived by triangulation. A clear example of 2,5D surface is the digital terrain models (DTM).

- Surfaces described by a function  $f(X, Y, Z) = 0$  are *3D surfaces*, namely closed surfaces characterized by holes and occlusion. The digital surface models (DSM) are 3D models.

*Triangle-* and *grid-based surface modeling* are the two most widely used approaches for surface reconstruction from a sparse *unstructured point cloud*. The first approach, however, is widely used for 3D surfaces modeling for its simplicity and versatility: the process leads to the formation of a Triangular Irregular Network (TIN) taking advantage of a wide variety of algorithms. Here Delaunay triangulation is discussed, as it is one of the most used. Moreover, since this chapter focuses on close-range photogrammetry technique, that mainly leads to reproduce medium and small-sized objects, the Delaunay triangulation technique for the 3D surfaces reconstruction is deepened.

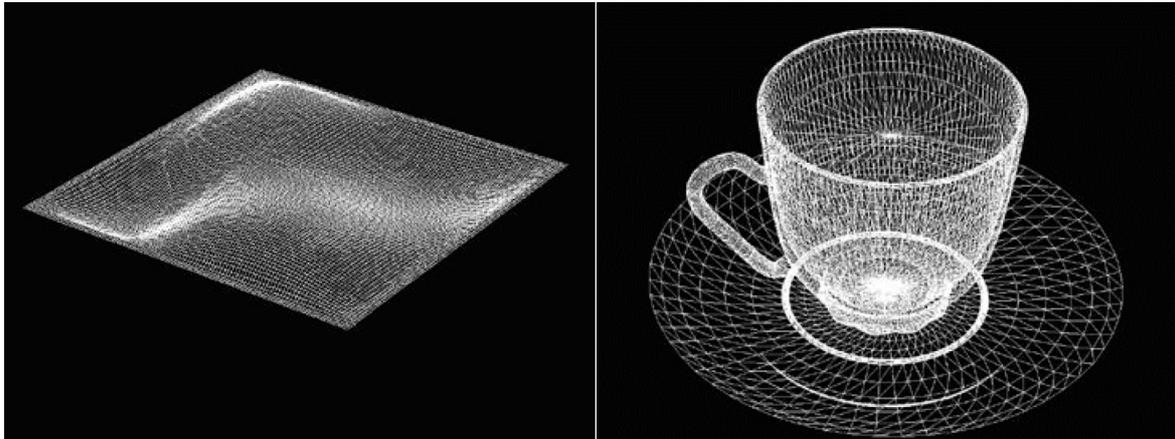


Figure 2.12. 2,5D surface on the left and 3D surface on the right. (Luhmann et al., 2019).

### Meshing: Delaunay triangulation

The simplest and often adopted method for the 3D surface model reconstruction from irregularly distributed data is the *triangle meshing*. Differently from a grid-based surface modeling, which uses polygonal surface elements, this approach takes advantage of the property of the triangular elements to guarantee the planarity of surface elements (that is the reason for the great employment of this 3D surface modeling). In fact, as anticipated above, the method combines three adjacent points of the cloud to form a surface element of triangular shape, which can be properly defined as a plane of parametric equation:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

The formation of a TIN requires in general three conditions to be satisfied:

1. The unicity of the TIN for a given data set if the same algorithm is used;
2. The minimization of the sum of the three edges of the triangles;
3. The local equiangularity.

This last condition ensures that the TIN is optimum, which is to say that the minimum angle is maximized and the maximum angle is minimized for all triangle (MAX-MIN angle principle). This condition is satisfied if, swapping the diagonals of the polygons formed by two adjacent triangles, the

interior angles do not decrease below their minimum values and do not increase above their maximum (Figure 2.13). The triangles must also satisfy the condition that any other points would not be included in the circumcircles of each triangle (Figure 2.14) (Li et al., 2005).

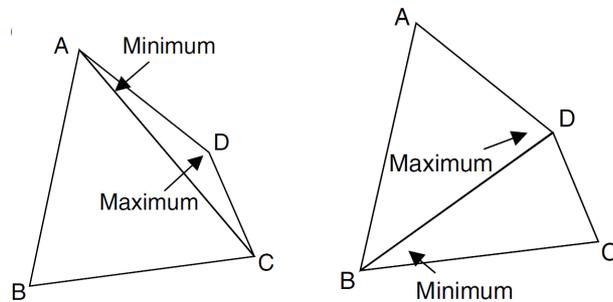


Figure 2.13 MIN-MAX angle principle for local equiangularity. On the left, before swapping the diagonal and on the right after the swapping (Li et al., 2005).

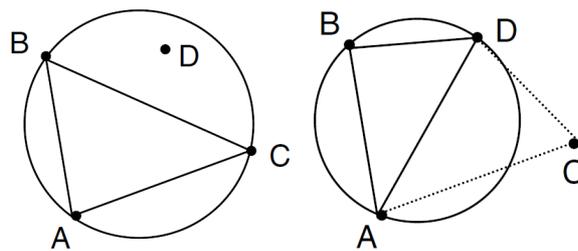


Figure. 2.14. Empty circumcircle principle. Since the circumcircle contains point D (left), it is used to form the ADB triangle that now satisfies the empty circumcircle condition (Li et al., 2005).

The mentioned requirements are well satisfied by Delaunay triangulation, which leads to a series of linked and non-overlapping triangles independent of the starting point. TIN can be built in static or dynamic approach, one that form a global network considering all the points (*static* approach) and the other one that progressively adds and/or removes points into the network during the triangulation (*dynamic* approach). When dealing with a huge amount of data, usually the dynamic approach is preferred for its capability to manage the amount of point gradually. The most practical dynamic triangulation algorithm is the Bowyer–Watson algorithm (also called simple *incremental algorithm*): it starts generating two adjacent coarse triangles and continues by adding points inside each triangle such that each added point become the vertex of three new enclosed triangles. At each step, the empty circumcircle principle must be satisfied. Figure 2.15 shows the process.

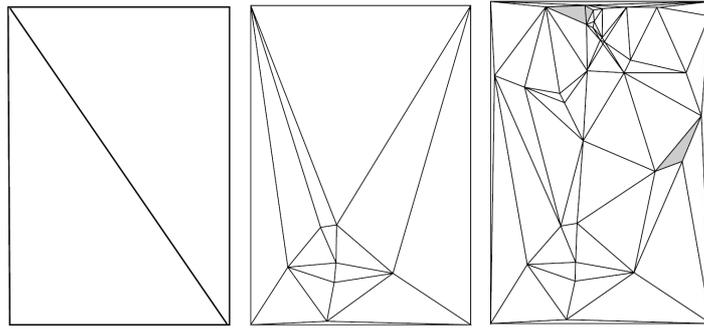


Figure 2.15. Dynamic Delaunay triangulation process. From left to right: two coarse triangles, intermediate step after the addition of points and final TIN (Li et al., 2005).

A triangular mesh is usually stored in triplets of X, Y, Z values that represents the coordinates of the three vertices of each triangle. This is the case of STL (stereo lithography) format, that stores the information creating a block for each triangle containing the normal vector and the vertices coordinates. Mainly used for rapid prototyping, 3D printing and computer-aided manufacturing, it does not contain any information about the texture properties. Since STL format stores the information redundantly (the vertices shared by neighbor triangles are duplicates), binary format sometimes is preferred for its efficient use of memory.

## 2.6 Multi-image acquisition

As introduced in previous paragraphs, high accuracy close-range photogrammetry is today often realized using the SfM process, that enables the automatic reconstruction of a 3D structure from a set of images. The most common data acquisition methodology in close-range photogrammetry is the *multi-view* (or *multi-photo*) approach (Figure 2.16): the scene is acquired by a camera moving around the object; the camera (nowadays often non-metric) acquires an unlimited number of images from different positions chosen to avoid any lack of data, namely to ensure sufficient intersecting angles of bundles of rays (Luhmann et al., 2019). This approach is suitable for close-range photogrammetry when the object shape has peculiarities, such as occlusions or hidden facets that can be acquired only from several different views, and/or when 3D model high-accuracy is required. A key point of the multi-view approach therefore, is to find the best configuration that satisfy the modelling requirement as resolution, precision and completeness. The selection of the camera and the image stations network are hence fundamental, as well as finding the optimal focal length and the distance from the object and between stations. In the next paragraphs are illustrated some generic rules for an optimal multi-image acquisition, that are valid for each scale object and for different platform.

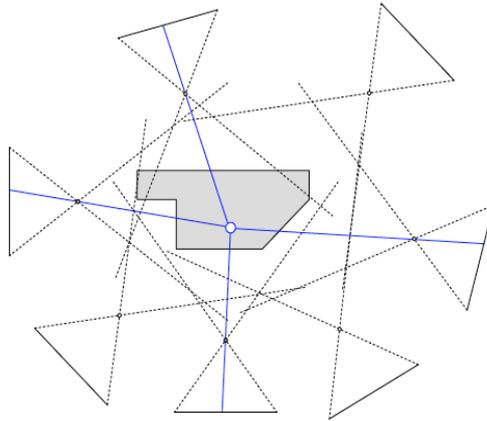


Figure 2.16. Multi-view configuration.

### 2.6.1 Geometric rules

Before the image acquisition of a certain scene, a good planning is fundamental to achieve the desired level of precision and resolution. The relation that characterize a stereo measurement in the stereo normal case is the following (Kraus, 2007):

$$d = -\frac{fB}{z_p} \quad [2.13]$$

Where:

$d$  is the disparity

$f$  is the focal length

$B$  is the baseline

$Z$  is the distance

$p$  is the pixel pitch

From [2.13] it is possible to define two components which represent respectively the image scale and the intersection angle.

$$f/Z = \text{image scale}$$

$$B/Z = \text{intersection angle}$$

The idea is to adopt the optimal capturing network configuration ensuring completeness and accuracy. In this phase it is therefore necessary to find the suitable intersection angle and image scale that ensure optimal matching performances and depth precision. In general, the following rules are valid:

- The point precision  $\sigma_p$  decreases with increasing distance  $D$  of the camera from the object;
- Small intersection angles lead to noisy results;
- Large baselines lead to lower image similarity, and consequently to scarce matching performance;

In practice, it is preferable to adopt small intersection angles to acquire additional models to be used for surface reconstruction, as well as to acquire highly overlapping imagery (high redundancy) that improve the object precision in the space. An important relation associated with the choice of the baseline regards the base-to-height ratio and the angle between surface normal and viewing ray (Figure 2.17). Here the concept of disparity is applied, that refers to the corresponding pixels' location difference between images. Non-flat surfaces (low base/height ratio, from 0 to approximately 0.2) accept high tilt angles (up to  $70^\circ$ ) maintaining low disparity shift, namely a good image matching. If the disparity shift become greater than 1 pixel, the matching could fail (Wenzel et al., 2013).

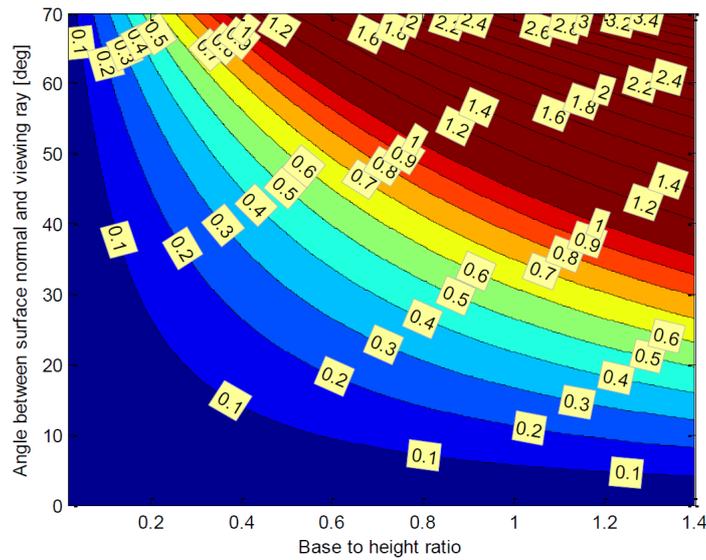


Figure 2.17. Disparity shift (yellow numbers) related to base-to-height ratio and angle between surface normal and viewing ray (Wenzel et al. 2013).

### 2.6.2 Acquisition procedure: 3x3 rules

Since close-range photogrammetry began to make practical use of the non-metric cameras, it has been necessary setting an acquisition strategy ensuring a metrically correct dense surface reconstruction. Waldhäusl and Ogleby (Waldhäusl and Ogleby, 1994) in 1994 introduced few simple rules for close-range photogrammetric documentation of architectural objects: they are known as CIPA 3x3 rules and give recommendations on the selection of the camera stations, suitable stereo models, camera and lens, according to the surface reconstruction requirement.

The strategy requires three steps, each of which is composed by other three sub-steps. Even though enunciated for the architectural documentation, they have a general valence in the field of high-accuracy close-range photogrammetry. Table 2.1 summarizes the 3x3 rules.

3 GEOMETRICAL RULES	3 PHOTOGRAPHIC RULES	ORGANIZATIONAL RULES
<ul style="list-style-type: none"> <li>▪ <b>Control information</b></li> <li>▪ Long distances between well-defined points (targets)</li> <li>▪ Plumb line</li> </ul>	<ul style="list-style-type: none"> <li>▪ <b>Camera geometry</b></li> <li>▪ No zooming</li> <li>▪ No shift optics</li> <li>▪ No distance changes</li> <li>▪ The true documents are the original negative or digital “RAW” equivalents.</li> <li>▪ Set the camera to use its highest quality format</li> </ul>	<ul style="list-style-type: none"> <li>▪ <b>Sketches</b></li> <li>▪ Draw the ground plan with the direction of the North</li> <li>▪ Draw location of control points with distances and plumb-line</li> <li>▪ Draw the location of natural points used</li> </ul>
<ul style="list-style-type: none"> <li>▪ <b>Multi-photo all-around coverage</b></li> <li>▪ Take a "ring" of pictures all around the object, overlapping each other more than 50%</li> <li>▪ Take shots from half the object’s height</li> <li>▪ Include parts of the neighbourhood</li> <li>▪ Add diagonal shots of two sides</li> <li>▪ Add traversing shots of the neighbourhood</li> <li>▪ Check multiple coverage</li> <li>▪ Add orthogonal full façade shots for overview and rectification</li> </ul>	<ul style="list-style-type: none"> <li>▪ <b>Camera calibration</b></li> <li>▪ Medium format</li> <li>▪ Large sensor</li> <li>▪ Wide-angle lenses are better than narrow-angle lens (very wide-angle lenses should be avoided)</li> <li>▪ Calibrate camera with fixed focus lens</li> <li>▪ Use standard colour chart</li> </ul>	<ul style="list-style-type: none"> <li>▪ <b>Write protocols</b></li> <li>▪ Site name, location, and georeference</li> <li>▪ Date, the weather and the personnel.</li> <li>▪ Camera, optics, focus and distance settings</li> <li>▪ Calibration report</li> <li>▪ Description of the place</li> </ul>
<ul style="list-style-type: none"> <li>▪ <b>Stereo partners</b></li> <li>▪ Normal case (base/distance = 1: 4 to 1: 5)</li> <li>▪ Convergent case (base/distance = 1: 10 to 1: 15)</li> <li>▪ Avoid divergent case</li> <li>▪ Add close-up stereo pairs for details</li> <li>▪ Check stereo coverage</li> <li>▪ Add more shots and take manually measures</li> </ul>	<ul style="list-style-type: none"> <li>▪ <b>Image exposure</b></li> <li>▪ Work with consistent exposure</li> <li>▪ Use HDR to capture unbalanced exposures</li> <li>▪ Use tripods and/or remote control to avoid camera movements</li> <li>▪ Use panoramic tripods head to get parallaxfree panoramic imagery</li> <li>▪ Black and white is sufficient but colours help the image matching</li> <li>▪ Geotagging is preferable</li> </ul>	<ul style="list-style-type: none"> <li>▪ <b>Final check</b></li> <li>▪ Completeness, accessibility and safety of data</li> <li>▪ Save RAW format</li> <li>▪ Do not crop any of the images, use the full format</li> </ul>

Table 2.1 adaptation of 3x3 rules introduced by Waldhäusl and Ogleby

Within the geometrical rules, once set the precision, the image scale is selected considering the limits represented by focal length and camera stations. The distance  $Z$  can be calculated through the following relation:

$$Z = \sqrt{fB \frac{\sigma_P}{\sigma_I}}$$

Where

$\sigma_P$  is the precision in object space

$\sigma_I$  is the precision in image space

The choice of camera, lens and camera settings is crucial for a good close-range photogrammetric survey. As Table 2.1 shows in the section “3 photographic rules”, large calibrated sensor cameras are advisable, as well as choosing a wide-angle lens and short fixed focal length, that ensure an efficient acquisition and a better image orientation. To ensure sharpness for the whole scene, which is necessary for a satisfactory feature extraction phase, it is good practice setting long exposure time and small aperture, this last ensuring sufficient depth of field. In general, underexposed images should be preferred since they are characterized by higher contrast than overexposed image, and ISO should be as low as possible. For these purposes, a tripod could be a simple solution for optimal image quality (Grussenmeyer et al., 2002; Wenzel et al., 2013).

Following the described set of rules, therefore, is the first step for a precise and complete surface documentation and modelling. Correct camera geometry and survey configuration is fundamental during the acquisition phase, as well as a good image exposure affects positively the photogrammetric image measurement (section 2.4). However, for satisfying results, cannot be taken out of consideration the image orientation and the effects that it has on image matching and 3D surface reconstruction. High accuracy close-range applications need high quality orientation and as explained before, the most commonly used approach for the automatic 3D structure reconstruction is the SfM approach.

As Figure 2.7 shows, SfM approach uses the bundle adjustment method for both interior and exterior parameters calculation (section 2.3). Bundle adjustment process with added Ground Control Points (GCP) is proven to be a good strategy that leads to higher accuracy of the final results and reduced scaling and registration errors (Oniga et al., 2018). On the other hand, the use of tie points, regardless of their number and distribution, do not change significantly the accuracy of object space (Khoshelham et al., 2012) and moreover, it can be reduced by using tie points with subpixel accuracy. However, using appropriate matching methods for tie points with good precision improves the accuracy of the bundle.

Next paragraph exposes a case study that enable to make some considerations about the procedures for a high accuracy documentation and modeling mentioned above.

## 2.7 Quality assessment of a close-range photogrammetric process: a case study

This study analyses the photogrammetric reconstruction processes taking in consideration several different parameters that should be accurately set during the survey. The aim is to define some good-practices for a close-range photogrammetric survey, starting from the multi-view configuration to the orientation strategy, as well as understanding how differently combined operations can affect the quality of the model. Wrong photogrammetric planning in fact, can lead to deformation of the 3D model, also described as bending effects, whose magnitude can reach values ten times larger than GSD. This is particularly valid in case of long flat objects, which require a sequence of photographs organized in strips, the so called open sequences (Nocerino et al., 2014). The case study is a brick-wall of dimensions 8x2,5m (Figure 2.18) and can be considered an architectural object for which Waldhäusl and Ogleby recommendations apply; therefore, the 3x3 CIPA rules have been followed, paying attention to geometrical and photographic rules.



Figure 2.18. The brick wall taken as case study for the photogrammetric process quality assessment.

The survey has been conducted with a non-metric full frame camera, Canon Eos 6D, with a CMOS sensor of 20.2 MP (36x24 mm), EF 24-70mm f/2.8 IS USM.

<i>Camera model</i>	<i>Resolution</i>	<i>Focal length</i>	<i>Pixel size</i>
Canon EOS 6D	5472 x 3648	50 mm	6.66 x 6.66 $\mu\text{m}$

The software used for the 3D multi-view reconstruction of the wall is *Agisoft Metashape*, one of the most employed image-based 3D modelling solution for the creation of 3D products from images. *Agisoft Metashape* realizes the SfM approach including the bundle adjustment solution (paragraph 2.5.1), and basically works following four main steps: (I) camera alignment, (II) point cloud generation, (III) Mesh/DEM reconstruction and (VI) texturing/orthomosaic generation. Since this

study aims to the evaluation of the quality of the whole photogrammetric process, some precise aspects of the workflow have been analysed in deep and in particular, two fundamental aspect of the photogrammetric process have been taken into consideration for the analysis:

- The multi-view camera configuration (paragraph 2.7.1)
- The photo orientation optimization (paragraph 2.7.2)

It was chosen to employ a series of target (Figure 2.19) and use them as known markers during the photogrammetric reconstruction. They have been attached on the surface of the test wall as Figure 2.20 shows.

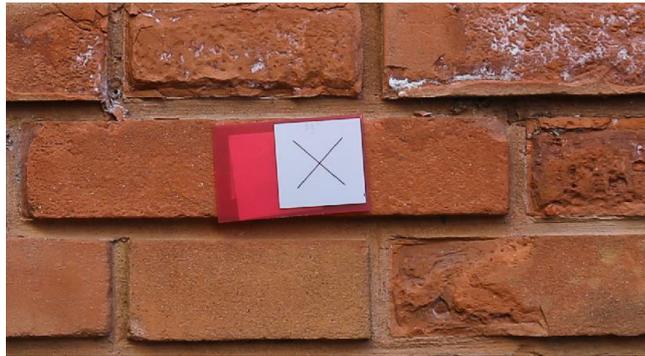


Figure 2.19. A target used for the photogrammetric survey.

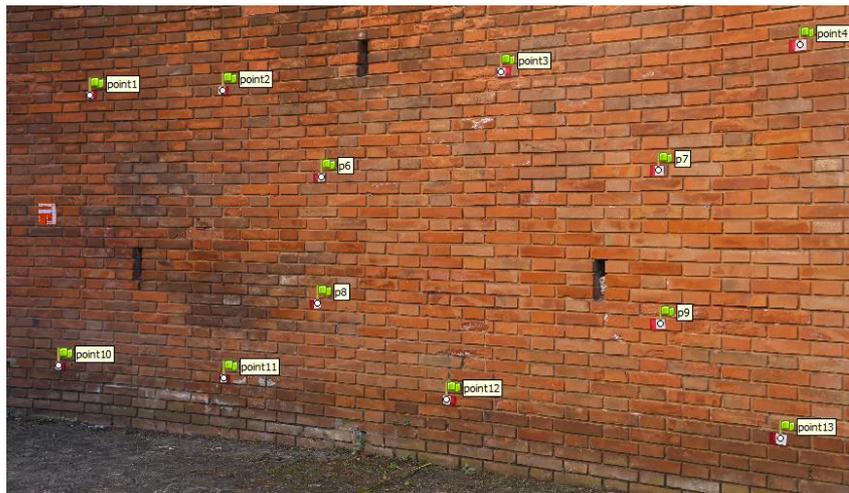


Figure 2.20. Placement of 14 targets on the test wall.

Before the image acquisition, the 14 targets were collimated through a first grade total station (Leica TS30) (Figure 2.21). This preliminary step is necessary to have available some coordinates of known points to be used as GCP. Table 2.2 contains the coordinates of the 14 known points.



Figure 2.21. Leica TS30 setting. In the background, the test wall with 14 targets.

	<i>X</i>	<i>Y</i>	<i>Z</i>
<b><i>P1</i></b>	6.76742	24.92113	2.54227
<b><i>P2</i></b>	7.86211	25.07491	2.53112
<b><i>P3</i></b>	9.91003	25.17704	2.52613
<b><i>P4</i></b>	11.60304	25.08527	2.51978
<b><i>P5</i></b>	13.07257	24.87590	2.52682
<b><i>P6</i></b>	8.64510	25.14229	1.95314
<b><i>P7</i></b>	10.87452	25.14947	1.94983
<b><i>P8</i></b>	8.61565	25.14633	1.16622
<b><i>P9</i></b>	10.88738	25.15114	1.16119
<b><i>P10</i></b>	6.50910	24.88168	0.65420
<b><i>P11</i></b>	7.88326	25.08474	0.65275
<b><i>P12</i></b>	9.55909	25.18044	0.64807
<b><i>P13</i></b>	11.52297	25.09823	0.64542
<b><i>P14</i></b>	13.06927	24.87682	0.57288

Table 2.2. X,Y,Z coordinates of 14 targets

In order to scale the model, some distances between targets were measured and used for scale bars addition (Table 2.3).

<b><i>P6-P7</i></b>	2.229 m
<b><i>P6-P8</i></b>	0.786 m
<b><i>P6-P9</i></b>	2.378 m
<b><i>P7-P8</i></b>	2.392 m
<b><i>P7-P9</i></b>	0.788 m
<b><i>P8-P9</i></b>	2.274 m

Table 2.3. Six measured distances between targets

### 2.7.1 Multi-view camera configuration

Planning a good configuration of the camera stations is generally an optimum prerequisite for an accurate reconstruction of the scene. The images of the test wall have been taken following the 3x3 recommendation table and the advises suggested by the software developers (however these two set of rules mainly correspond). Therefore, the photos satisfy the minimum overlapping requirement and accomplish the high redundancy goal; they have been shot with a regular parallel stripe configuration at different heights, starting from an elevation of 0,50 m from the ground to the highest elevation of 2 m, where the targets 1, 2, 3, and 4 were placed. The distance from the wall has been maintained constant, as well as the step between the camera stations. Table 2.4 collects all the measurements that characterize the photoset.

	<i>Nadiral</i>	<i>Convergent</i>	<i>Full facade</i>
<b><i>Distance from the wall</i></b>	2 m	2 m	6 m
<b><i>Step between cameras</i></b>	0,50 m	0,50 m	4 m
<b><i>Horizontal stripes</i></b>	4	3	1
<b><i>Spacing between stripes</i></b>	0,50 m	0,70 m	-

Table 2.4. Nadiral photoset measurements

Three separate photo set have been acquired respecting these spatial constraints and varying the optical axis inclination with respect to the wall surface. Three image configurations result: nadiral/normal, convergent, and full façade. The reason for the separate acquisition of three sets, lies in the intention to carry out tests on 3D reconstruction quality according to the configuration change

and/or according to the possible sets combination. Therefore, the following configurations are considered:

- A. Nadiral
- B. Nadiral + full façade
- C. Convergent
- D. Convergent + full façade
- E. Nadiral + convergent
- F. Nadiral + convergent + full façade

Images acquisition required the use of a tripod to keep the height of the camera constant during the survey. Moreover, a measuring tape was used to maintain the distance from the wall. Totally, 100 images were taken (Table 2.5). Figure 2.22, 2.23, and 2.24 shows the three basic configurations of the camera positions.

<i>N° of images</i>	
<i>Nadiral</i>	56
<i>Convergent</i>	42
<i>Full façade</i>	2
<i>tot</i>	<i>100</i>

Table 2.5. Number of images

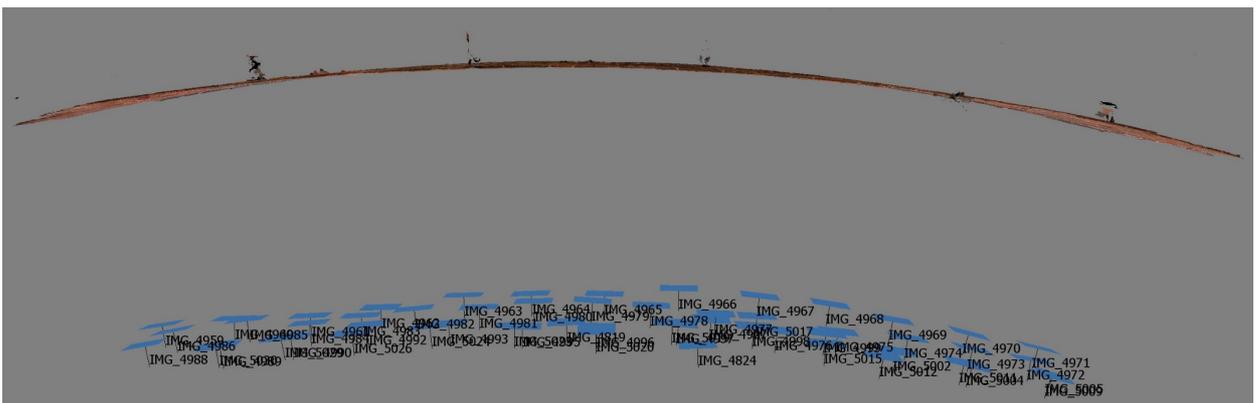


Figure 2.22. Nadiral images (in blue) with respect to the wall (top view).

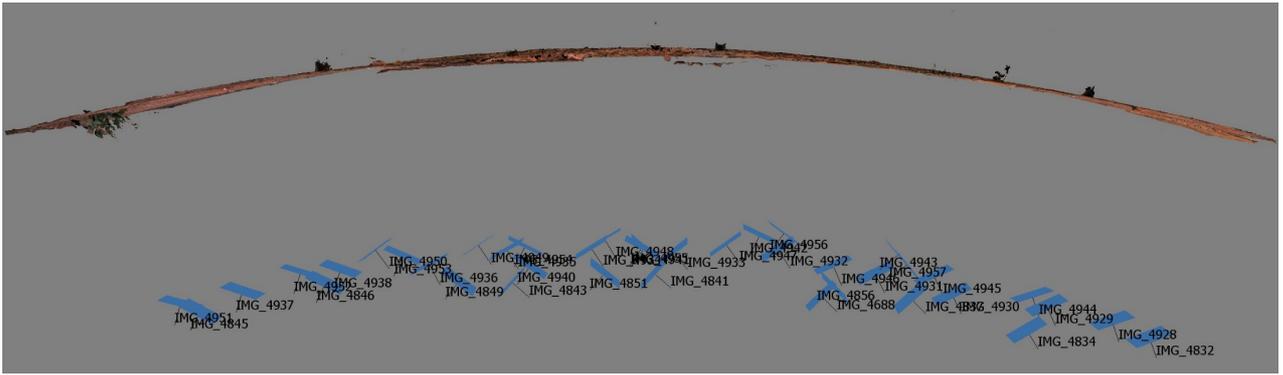


Figure 2.23. Convergent images (in blue) with respect to the wall (top view).

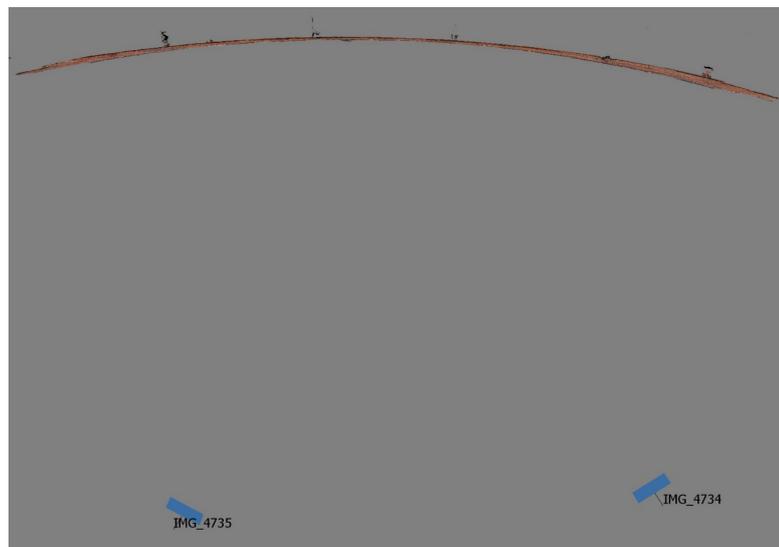


Figure 2.24. Full façade images (in blue) with respect to the wall (top view).

### 2.7.2 Photo orientation optimization

The photo orientation is the first step of the processing and Agisoft Metashape allows to perform it directly on the acquired photos. The procedure estimates the tie point positions taking advantage of feature-based matching algorithm. It is well known that the accuracy of the orientation can be improved placing markers and using them as tie points and/or GCP (if their spatial coordinates are known). Additionally, the application of a scale bar between markers of known distance, as well as ensuring a scaling of the model, can prove to be useful when the employment of GCP on the scene is not possible. Therefore, this part of the study aims to evaluate the quality of 3D reconstruction varying the photo orientation strategy based on the employment of the 14 targets and the scale bar. The following alignment approaches are considered:

1. Photo orientation using the 14 target as tie points and scale bars;
2. Photo orientation using 4 targets as GCP;
3. Photo orientation using 10 targets as tie points and 4 targets as GCP;
4. Photo orientation using 14 targets as GCP;

For the approach 1 (Figure 2.25), all the targets located on the wall are considered as tie points and the scale bars are used to scale the model. This photo orientation configuration simulates a quick photogrammetric survey where only the digital camera and a measuring tape are employed.

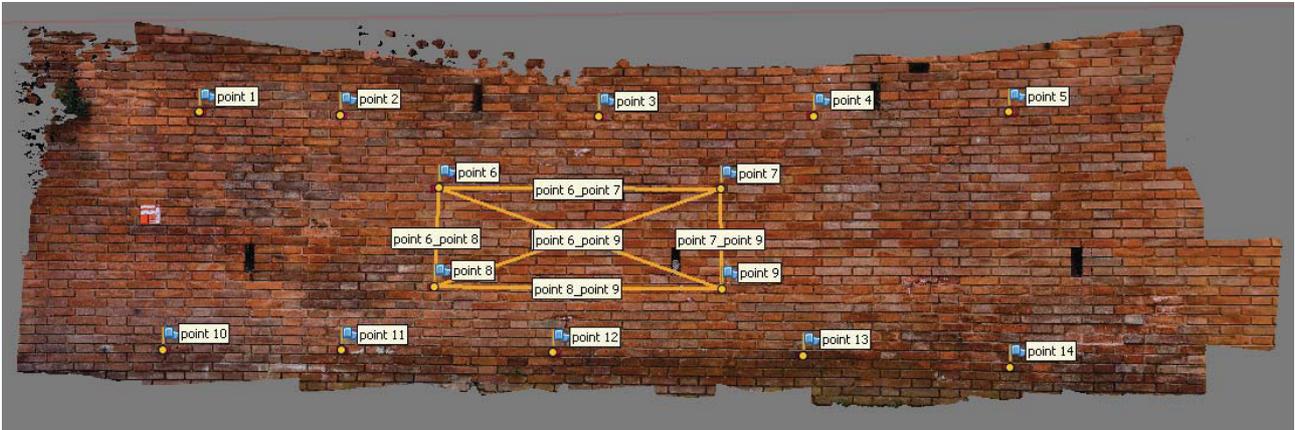


Figure 2.25. Approach 1: 14 targets as tie points and six scale bars (in yellow).

Approach 2 (Figure 2.26) involves the automatic scaling of the model through the coordinates of 4 targets: P6, P7, P8 and P9 are therefore considered as GCP. This orientation strategy aims to reproduce a photogrammetric survey less quick than the previous one, that employs a total station to collimate some points, which, however, are distributed non-uniformly on the scene.



Figure 2.26. Approach 2: 4 targets used as GCP. They are badly distributed since do not cover uniformly the wall but only the central part.

Approach 3 (Figure 2.27) can be considered as a variation of the approach 2, where 10 tie points are added to 4 GCPs. The tie points are comparable to natural points. The aim is to evaluate if how much the addition of tie points improves the model reconstruction accuracy.

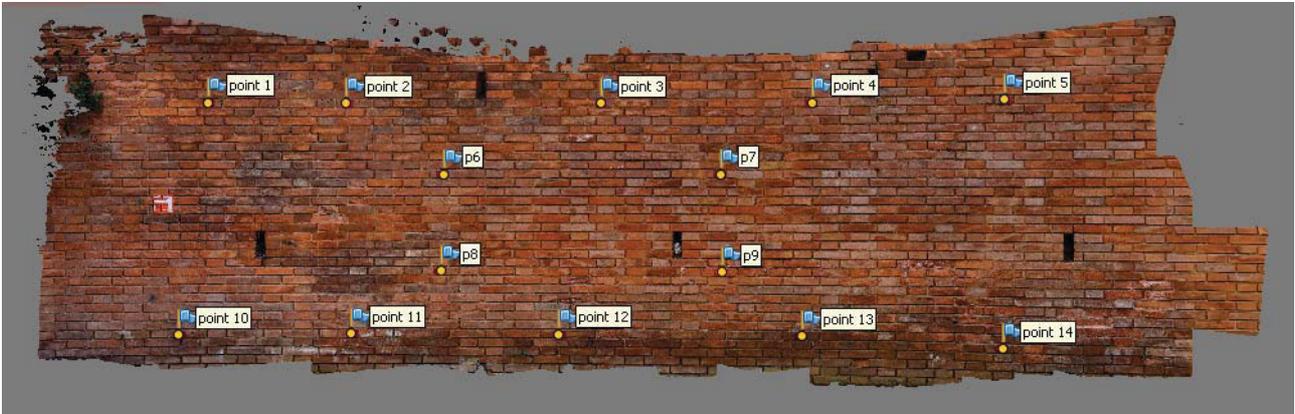


Figure 2.27. Approach 3: 10 targets are considered as tie points (point 1, point 2, point 3, point 4, point 5, point 10, point 11, point 12, point 13, point 14), while 4 central targets are used as GCPs.

Approach 4 (Figure 2.28) considers all targets as GPSs. Therefore, all the coordinates are loaded and used to tie the configuration during the photo orientation. This configuration should ensure the most precise result, even though not frequently applicable for the demanding and massive use of a total station.

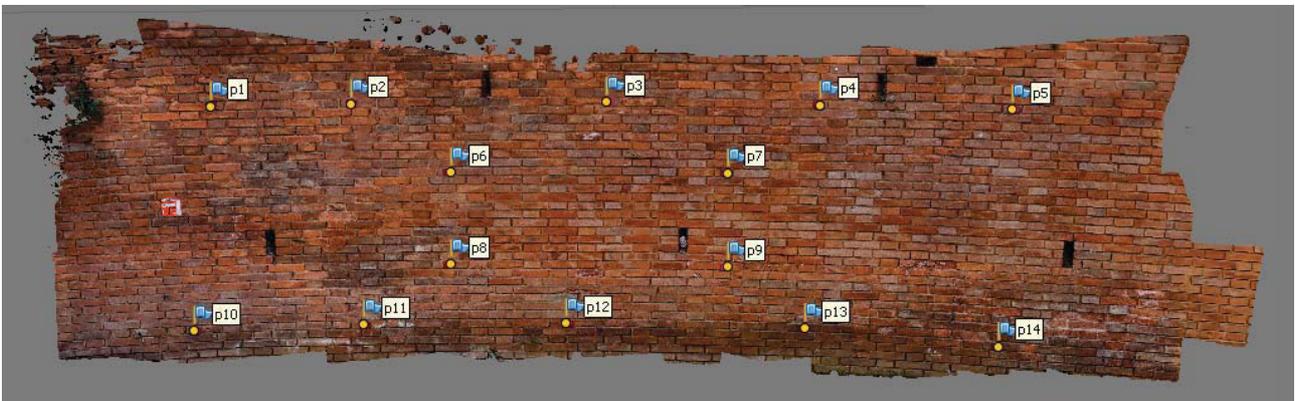


Figure 2.28. Approach 4: all the targets coordinates are used for the photo orientation. The 14 targets become 14 GCP.

### 2.7.3 Results and considerations

The quality assessment of the 3D reconstruction of the wall involves the comparison of the photogrammetric models varying survey configuration and photo orientation approaches, with another independent model, in this case reconstructed through a TLS Riegl VZ400. The last has been considered as reference model for the qualitative analysis. Figure 2.29 shows the TLS model, Table 2.6 summarizes its characteristics. The analysis is conducted specifically on the DEMs; the DEM generated from the laser scanner point cloud has been taken as reference (DEM Riegl), and the DEMs generated from the photogrammetric reconstructions (one for each combination of parameters) have

been compared to it. The comparisons have been performed in QGIS environment, that allow to perform mathematical operations between DEMs.



Figure 2.29. Riegl VZ400 point cloud.

<i>N° TLS stations</i>	1
<i>Resolution of the model</i>	1,5mm
<i>N° points</i>	6.342.015

Table. 2.6. Characteristics of the Riegl VZ400 survey and model.

The first considerations regard the analysis of DEMs varying the photo orientation approach (1,2,3,4) at fixed camera configuration (A,B,C,D,E,F). The models have been compared in terms of *difference* between DEMs: white colour indicates a variation in the range of -3mm and +3mm, blue colour between -10mm and -5 mm, and red colour between +5.1 and +10 mm. The results follow.

### **Configuration A (Nadiral images):**

The photogrammetric reconstruction that involves nadiral image only, results to ensure a good quality models, at varying photo orientation strategy: in all the cases, the RMS reprojection error is lower than 0,3 pixels, and the photo orientation time is lower than 2 minutes. Table 2.7 shows the values of these parameters for each photo orientation strategy. Interesting considerations about the GCPs error estimation can be made: as Table 2.8 shows, the RMSE of GCPs in orientation mode 4 is more than twice the RMSE in configuration 2 and 3; however, the value is proved to be influenced by the GCP located in the extremities of the wall (Figure 2.30), as expected from literature (Nocerino et al., 2014). For the same reason, as comes up from the comparison with the Riegl DEM (Figure 2.31), the orientation approach 4 generates an significant deformation at the extremities of the model, that result to be warped, within a difference that reaches values up to 10 mm. On the other hand, the orientation approaches 1, 2, and 3, provide a moderate deformation, showing a slight distortion in the

low-central part. The final 3D model characteristics are summarized in Table 2.9. For all the photo orientation approaches, the degree of resolution of the 3D models is satisfying.

**Camera configuration A – Photo alignment quality**

	<b><i>RMS reprojection error</i></b>	<b><i>Maximum reprojection error</i></b>	<b><i>Matching and alignment time</i></b>
<b><i>Approach 1</i></b>	0.264618 pix	10.6969 pix	1 min 47 sec
<b><i>Approach 2</i></b>	0.26277 pix	10.7059 pix	1 min 42 sec
<b><i>Approach 3</i></b>	0.26425 pix	10.6959 pix	1 min 47 sec
<b><i>Approach 4</i></b>	0.271684 pix	10.5918 pix	1 min 47 sec

Table 2.7. Comparison between significant parameters for the photo orientation quality estimation varying the photo orientation approaches.

**Camera configuration A – Ground Control Points RMSE (mm)**

	<b><i>X</i></b>	<b><i>Y</i></b>	<b><i>Z</i></b>	<b><i>XY</i></b>	<b><i>Total</i></b>
<b><i>Approach 2</i></b>	0.676018	0.20886	0.89147	0.707548	1.3813
<b><i>Approach 3</i></b>	0.659201	0.213342	0.906064	0.692865	1.14062
<b><i>Approach 4</i></b>	1.76868	1.10062	2.18969	2.08317	3.0223

Table 2.8. Ground Control Points RMSE.

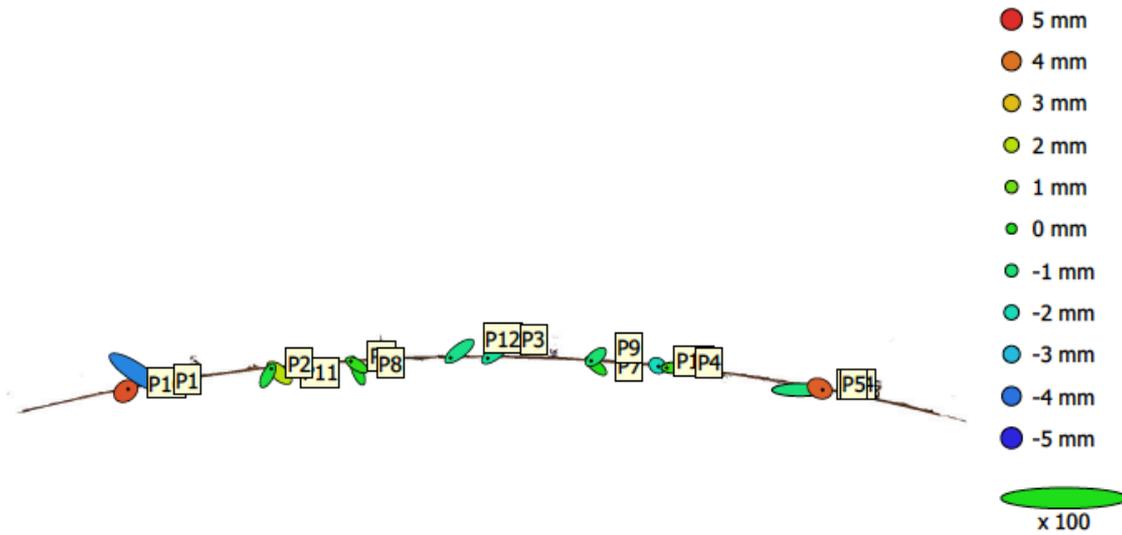


Figure 2.30. GCP location and estimates (approach 4). The ellipse colour represents the Z error, while the shape represents the X, Y errors.

	<i>Model face count</i>	<i>Ground resolution</i>
<b><i>Orientation 1</i></b>	14.010.115	0.281 mm/pix
<b><i>Orientation 2</i></b>	14.010.348	0.274 mm/pix
<b><i>Orientation 3</i></b>	14.010.543	0.274 mm/pix
<b><i>Orientation 4</i></b>	14.012.214	0.274 mm/pix

Table 2.9. Final 3D model resolution characteristics.

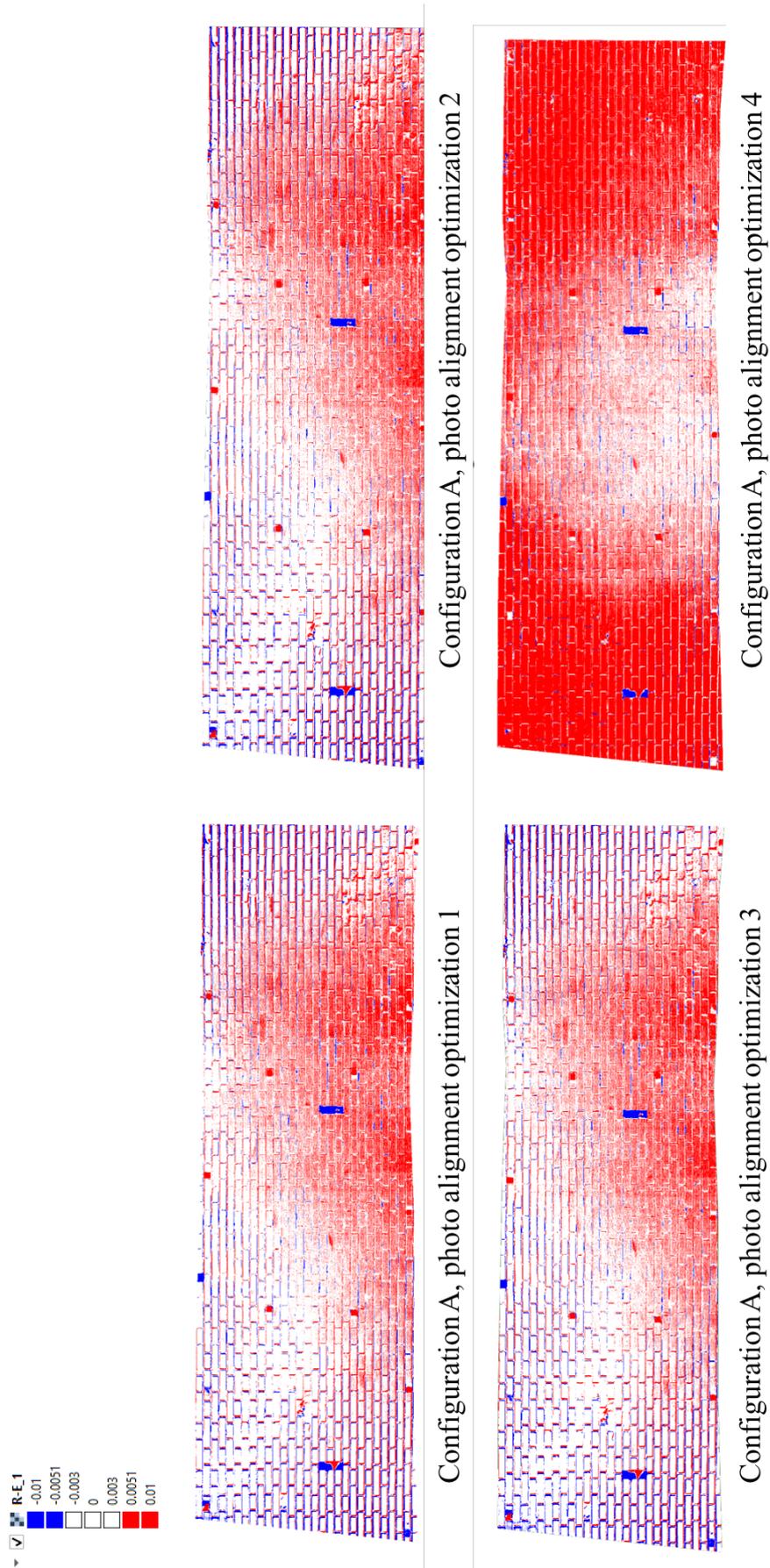


Figure 2.31. Deformation analysis of the models: comparison between DEM from the photogrammetric reconstruction and Riegl DEM.

**Configuration B (Nadiral + full façade):**

The configuration B differs from the previous one for the addition of two full façade images. The analysis of the reconstruction parameters shows however that the full façade images does not influence significantly the photo orientation optimization 2 and 4 with respect to the configuration A: the model deformation results for these two strategies (Figure 2.33) is essentially the same of the configuration A, both in terms of values of deformation and both in terms of deformations' distribution over the model surface. The photo orientation optimization 1 and 3 in presence of full façade images instead, are badly influenced: the models shows relevant deformations at the extremities, that in the strategy 3 are even higher and more spread than the discussed deformations of the approach 4. Analysing the orientation data however, can be noticed an increase in the reprojection error in all the cases (Table 2.10); GCPs errors result in general lower than in the previous case, maintaining however the same behaviour in the photo orientation approach 4 (Table 2.11, Figure 2.32). The final 3D model does not differ much from the one generated with camera configuration A (Table 2.12), maintaining a satisfying degree of ground resolution.

**Camera configuration B – Photo orientation quality**

	<i>RMS reprojection error</i>	<i>Maximum reprojection error</i>	<i>Matching and orientation time</i>
<i>Approach 1</i>	0.64513 pix	31.4685 pix	2 min 1 sec
<i>Approach 2</i>	0.620955 pix	32.1092 pix	1 min 55 sec
<i>Approach 3</i>	0.647464 pix	31.5587 pix	1 min 55sec
<i>Approach 4</i>	0.646519 pix	31.5408 pix	1 min 55 sec

Table 2.10. Comparison between significant parameters for the photo orientation quality estimation varying the photo orientation approaches.

**Camera configuration B- Ground Control Points RMSE (mm)**

	<i>X</i>	<i>Y</i>	<i>Z</i>	<i>XY</i>	<i>Total</i>
<i>Approach 2</i>	0.205678	0.231096	0.414851	0.309368	0.517504
<i>Approach 3</i>	0.132136	0.212689	0.25789	0.250393	0.359449
<i>Approach 4</i>	1.0727	1.30371	0.908983	1.68829	1.91744

Table 2.11. Ground Control Points RMSE.

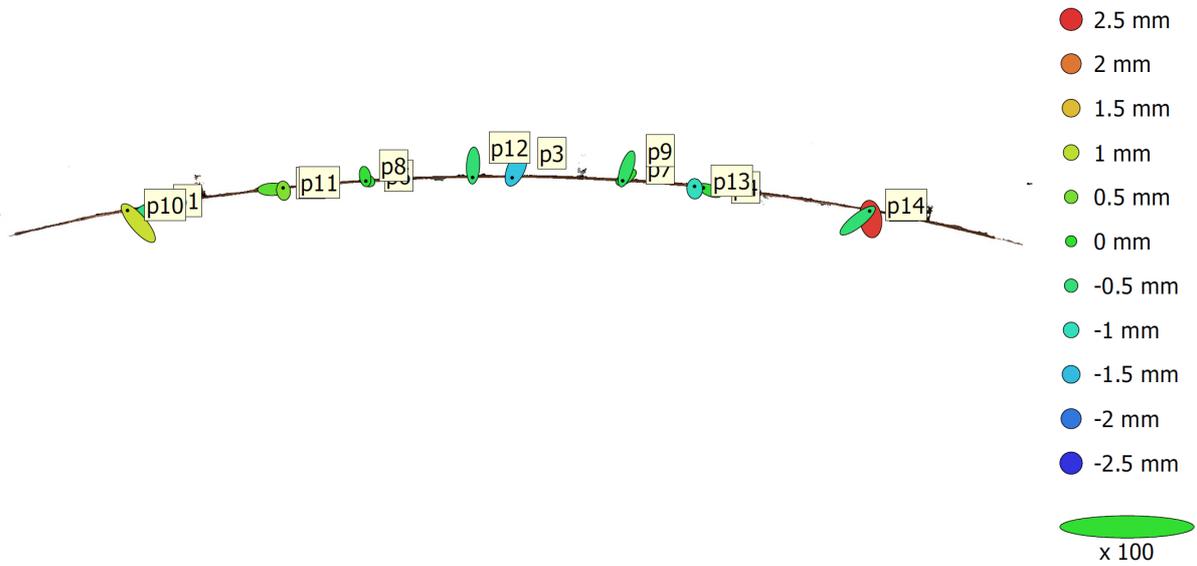


Figure 2.32. GCPs location and estimates (approach 4). The ellipse colour represents the Z error, while the shape represents the X, Y errors.

	<i>Model face count</i>	<i>Ground resolution</i>
<i>Approach 1</i>	14.028.325	0.301 mm/pix
<i>Approach 2</i>	14.022.899	0.274 mm/pix
<i>Approach 3</i>	14.032.715	0.274 mm/pix
<i>Approach 4</i>	14.036.973	0.274 mm/pix

Table 2.12. Final 3D model resolution characteristics

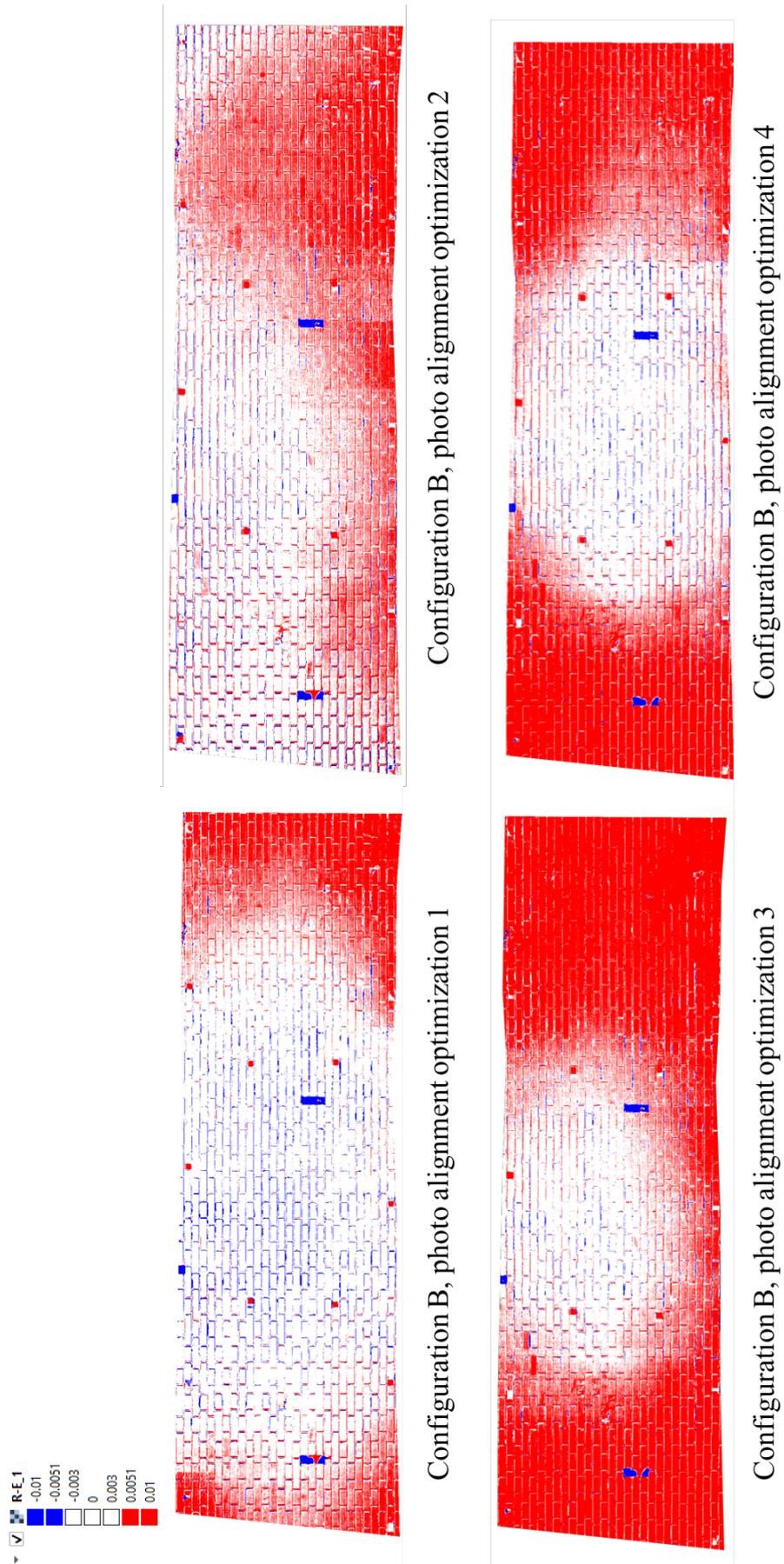


Figure 2.33. Deformation analysis of the models: comparison between DEM from the photogrammetric reconstruction and Riegl DEM.

**Configuration C (Convergent images):**

The acquisition phase with only convergent images results, as expected, inconvenient for both deformation analysis and photo orientation quality. The model generated using this camera configuration shows numerous anomalies for every photo orientation approach, that make the 3D model scarcely reliable. The reprojection errors are close to 1 pixel (Table 2.13) and the GCPs error reaches its maximum value in Z direction in correspondence to a GCP located at the left extremity of the wall (Figure 2.343); the RMSE of Ground Control Points reach the value of 5 mm using orientation approach 4 (Table 2.14). Analysing the deformations, any recurring element or behaviour can be spotted from the comparison between the Riegl DEMs and the photogrammetric DEMs, neither in approach 4, that shows a randomly distributed deformation pattern over the whole surface (Figure 2.35). The resolution of the final model is not undermined, and reaches values of around 0.3 mm/pix for all the approaches; the same consideration is valid for the face count, which slightly change varying the orientation strategy (Table 2.15).

**Camera configuration C – Photo orientation quality**

	<i>RMS reprojection error</i>	<i>Maximum reprojection error</i>	<i>Matching and orientation time</i>
<i>Approach 1</i>	0.876345 pix	34.4649 pix	1 min 28 sec
<i>Approach 2</i>	0.830349 pix	28.0981 pix	1 min 28 sec
<i>Approach 3</i>	0.877627 pix	34.8031 pix	1 min 28 sec
<i>Approach 4</i>	0.878366 pix	34.9860 pix	1 min 28 sec

Table 2.13. Comparison between significant parameters for the photo orientation quality estimation varying the photo orientation approaches.

**Camera configuration C- Ground Control Points RMSE (mm)**

	<i>X</i>	<i>Y</i>	<i>Z</i>	<i>XY</i>	<i>Total</i>
<i>Approach 2</i>	0.636522	0.309347	0.551515	0.707711	0.897232
<i>Approach 3</i>	0.788191	0.242032	0.580743	0.824515	1.008510
<i>Approach 4</i>	2.87135	2.30289	3.44761	3.68075	5.04321

Table 2.14. Ground Control Points RMSE.

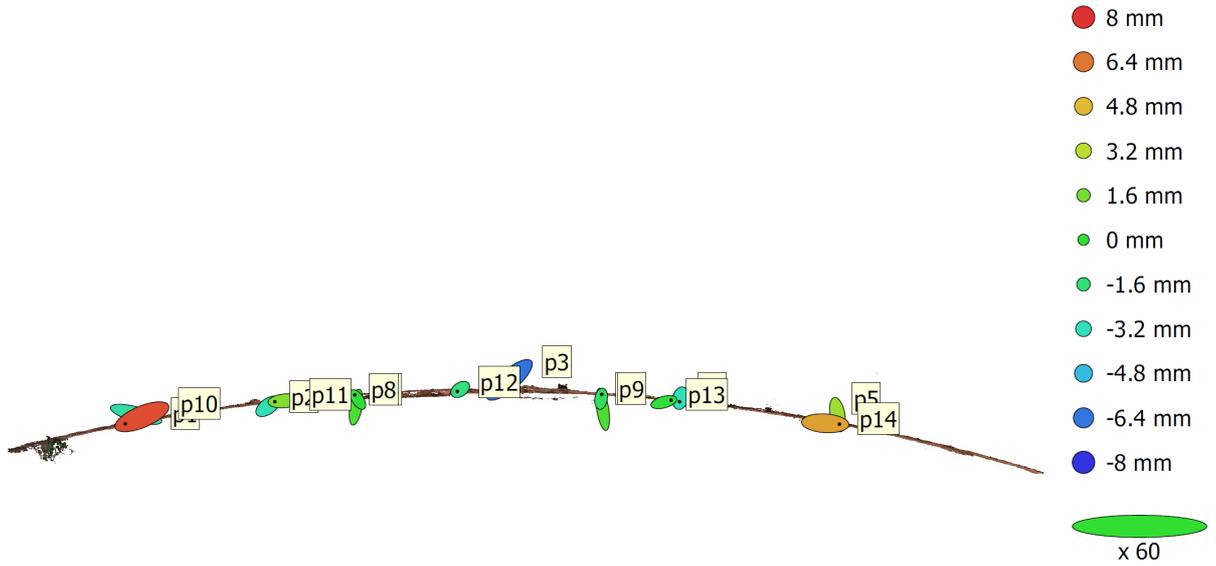


Figure 2.34. GCPs location and estimates (approach 4). The ellipse colour represents the Z error, while the shape represents the X, Y errors.

	<i>Model face count</i>	<i>Ground resolution</i>
<i>Approach 1</i>	14.297.124	0.325 mm/pix
<i>Approach 2</i>	14.119.308	0.317 mm/pix
<i>Approach 3</i>	14.300.084	0.318 mm/pix
<i>Approach 4</i>	14.300.003	0.318 mm/pix

Table 2.15. Final 3D model resolution characteristics

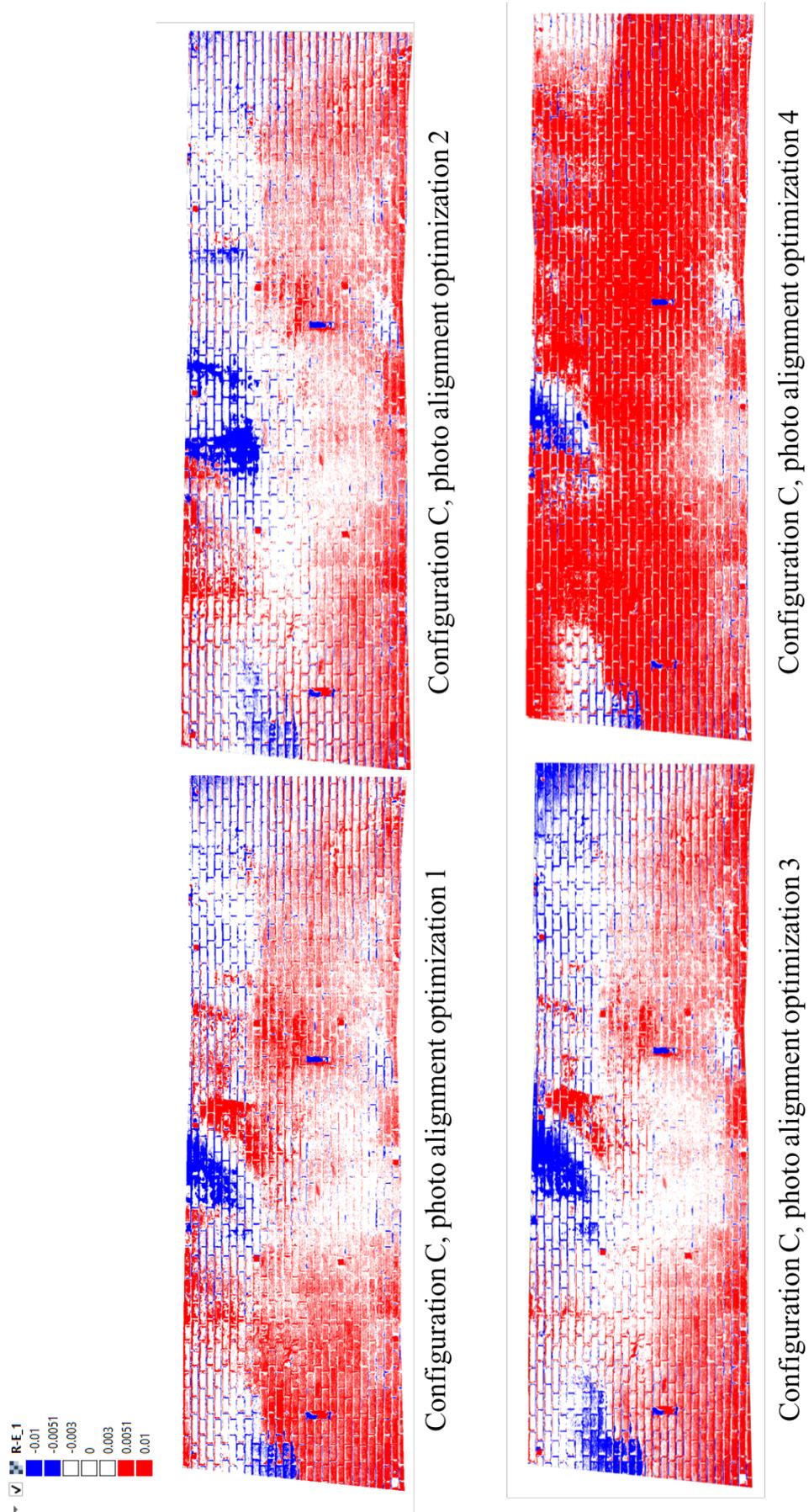


Figure 2.35. Deformation analysis of the models: comparison between DEM from the photogrammetric reconstruction and Riegl DEM.

**Configuration D (Convergent + full facade):**

Camera configuration D shows the same deficiencies as the configuration B in terms of models' deformation: the two full façade images do not change significantly the model surface reconstruction, resulting in almost the same randomly distributed envelope of configuration C (Figure 2.37). In terms of photo orientation quality, the presence of full façade images relatively influences the reprojection error and the GCPs RMSE, which are almost the same of the configuration C (Table 2.16 and 2.17). Figure 2.36 shows moreover a behaviour of GCPs error similar to the one noticed in configuration B. In general, the reprojection error is badly affected by the full façade images, reaching values greater than 1 pixel, while the processing time during the orientation procedure is still reduced and assume values under 2 min. The final models are characterized by high resolution, that differ slightly from the values reached in configuration C (Table 2.18).

**Camera configuration D – Photo orientation quality**

	<b><i>RMS reprojection error</i></b>	<b><i>Maximum reprojection error</i></b>	<b><i>Matching and orientation time</i></b>
<b><i>Approach 1</i></b>	1.24748 pix	34.9364 pix	1 min 36 sec
<b><i>Approach 2</i></b>	1.06431 pix	33.0045 pix	1 min 36 sec
<b><i>Approach 3</i></b>	1.29424 pix	35.8601 pix	1 min 36 sec
<b><i>Approach 4</i></b>	1.28913 pix	35.8713 pix	1 min 36 sec

Table 2.16. Comparison between significant parameters for the photo orientation quality estimation varying the photo orientation approaches.

**Camera configuration D- Ground Control Points RMSE (mm)**

	<b><i>X</i></b>	<b><i>Y</i></b>	<b><i>Z</i></b>	<b><i>XY</i></b>	<b><i>Total</i></b>
<b><i>Approach 2</i></b>	0.308695	0.211718	0.870565	0.374322	0.947629
<b><i>Approach 3</i></b>	0.533744	0.324274	1.507360	0.624529	1.631610
<b><i>Approach 4</i></b>	2.02018	3.5618	3.40358	4.09482	5.32465

Table 2.17. Ground Control Points RMSE.

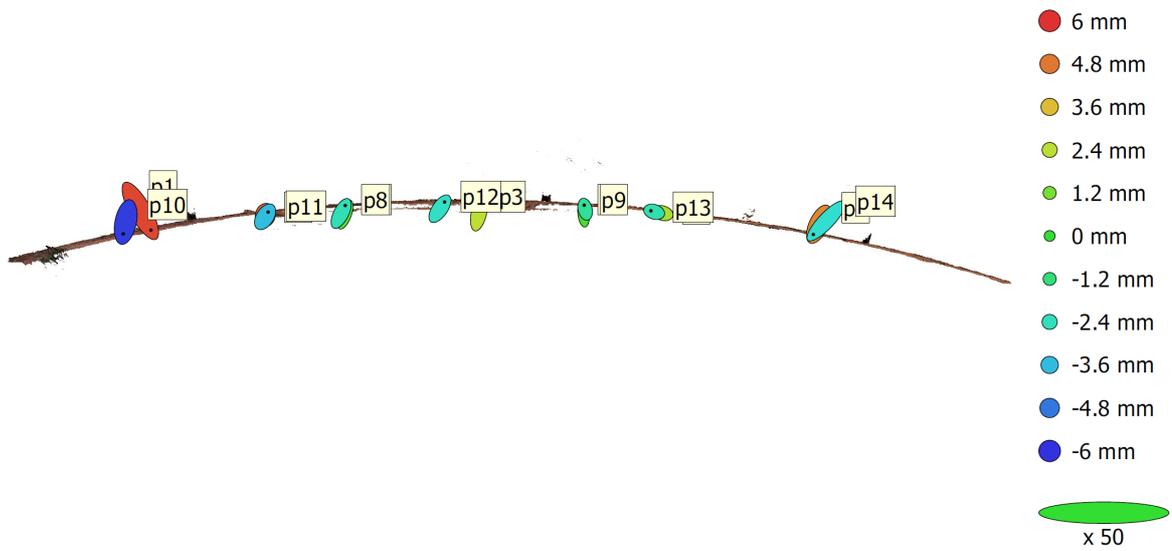


Figure 2.36. GCPs location and estimates (approach 4). The ellipse colour represents the Z error, while the shape represents the X, Y errors.

	<i>Model face count</i>	<i>Ground resolution</i>
<i>Approach 1</i>	14.198.276	0.353 mm/pix
<i>Approach 2</i>	14.257.561	0.318 mm/pix
<i>Approach 3</i>	14.242.642	0.318 mm/pix
<i>Approach 4</i>	14.248.168	0.318 mm/pix

Table 2.18. Final 3D model resolution characteristics

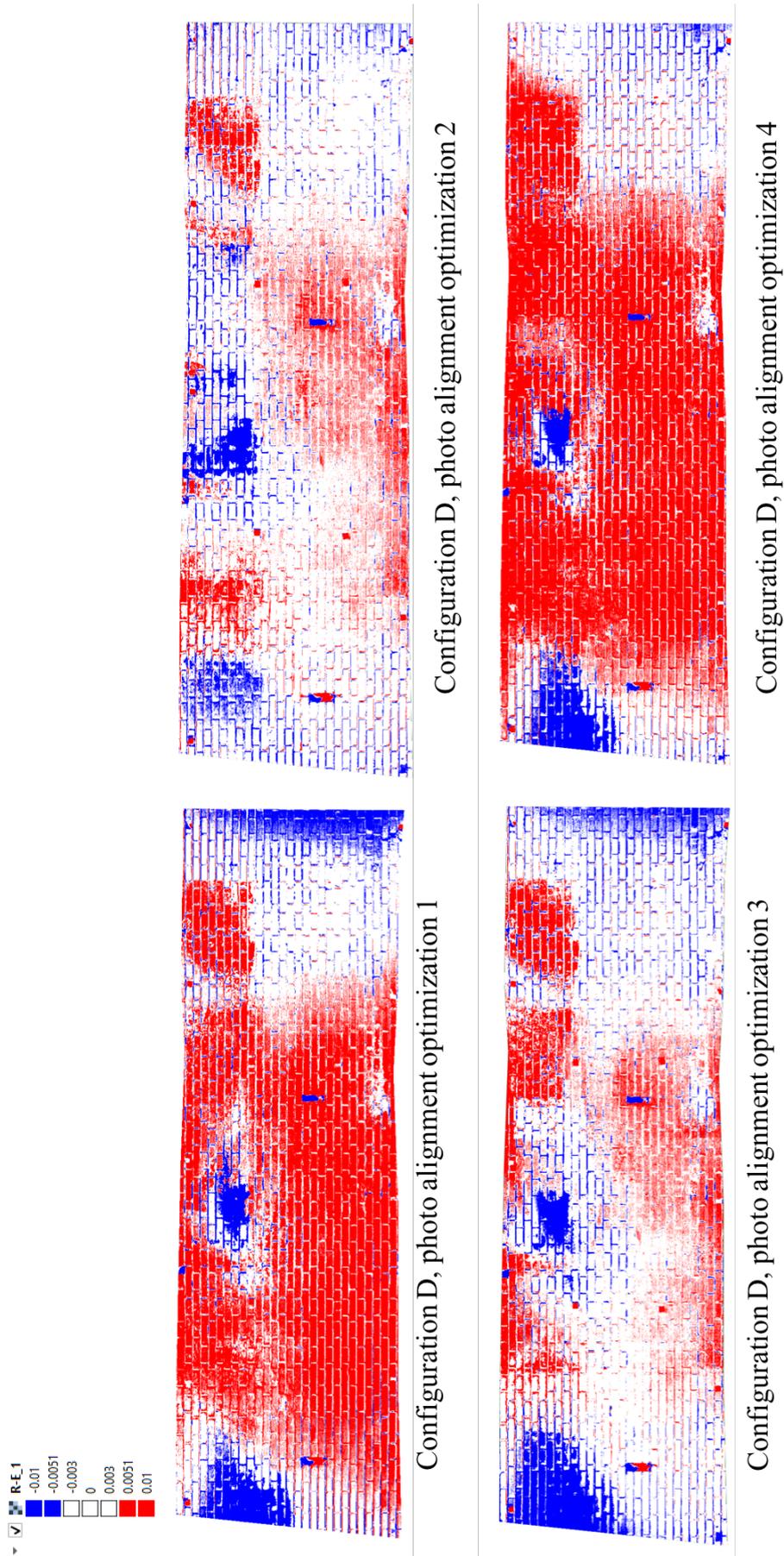


Figure 2.37. Deformation analysis of the models: comparison between DEM from the photogrammetric reconstruction and Riegli DEM.

**Configuration E (Nadiral + Convergent):**

This configuration, as expected, gives back more satisfactory results than the configuration C and D, thanks to the introduction of nadiral images that stabilize the photogrammetric reconstruction process. The models' deformations (Figure 2.38) are moderate in orientation 1, 2 and 3, even though the side effects related to the extremities of the objects remain. Orientation 4 has still numerous problems, regarding both the deformations and both the GCPs errors. Figure 2.39 shows that the model reconstructed with approach 4 presents a diffuse deformation within values in the range of 5-10 mm over more than the 50% of the surface. In Table 2.19 are reported the reprojection errors, that result to be near to 1 pix. The Ground Control Points RMSE (Table 2.20) are slightly lower than the previous case, but as Figure 2.33 shows, the location errors are maximum in correspondence of the extremities (respectively 8mm and 6.4 mm) in the Z direction. Despite deformation problems and scarce GCP accuracy in approach 4, the overall quality of the models in terms of resolution is confirmed (Table 2.21): in all the cases, the models' face count corresponds to 17.000.000 with a ground resolution slightly lower than 0.3 mm.

**Camera configuration E – Photo orientation quality**

	<b>RMS reprojection error</b>	<b>Maximum reprojection error</b>	<b>Matching and orientation time</b>
<b>Approach 1</b>	0.949184 pix	28.3171 pix	3 min 54 sec
<b>Approach 2</b>	0.849023 pix	27.69540 pix	3 min 54 sec
<b>Approach 3</b>	0.948728 pix	28.33460 pix	3 min 54 sec
<b>Approach 4</b>	0.950779 pix	28.58490 pix	3 min 54 sec

Table 2.19. Comparison between significant parameters for the photo orientation quality estimation varying the photo orientation approaches.

**Camera configuration E- Ground Control Points RMSE (mm)**

	<b>X</b>	<b>Y</b>	<b>Z</b>	<b>XY</b>	<b>Total</b>
<b>Approach 2</b>	0.24779	0.223107	0.731031	0.333431	0.803481
<b>Approach 3</b>	0.425471	0.25115	0.693054	0.494067	0.851132
<b>Approach 4</b>	2.92296	1.02998	3.15117	3.09912	4.41978

Table 2.20. Ground Control Points RMSE.

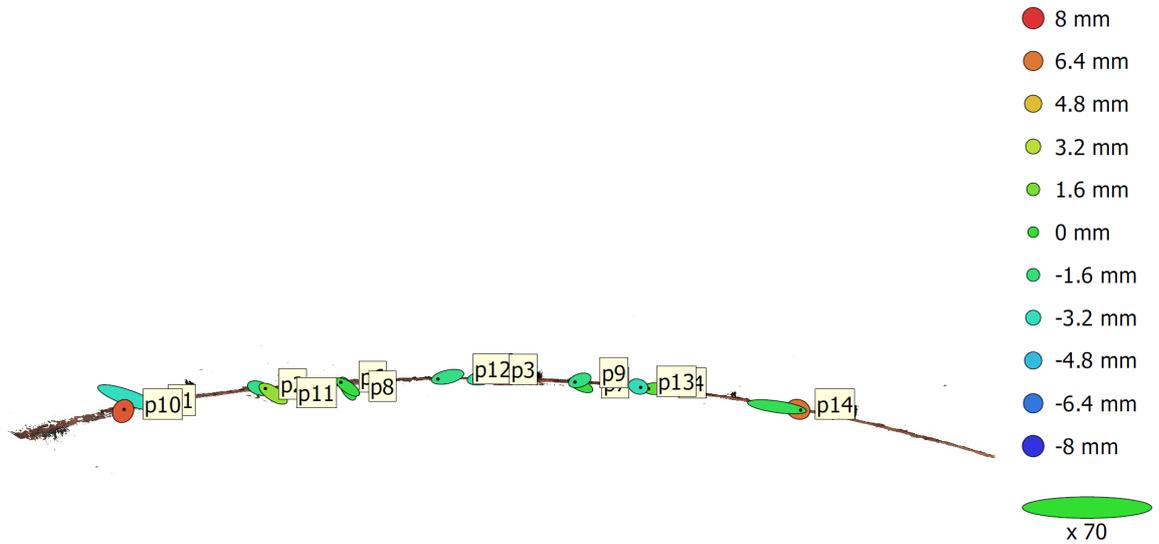


Figure 2.38. GCPs location and estimates (approach 4). The ellipse colour represents the Z error, while the shape represents the X, Y errors.

	<i>Model face count</i>	<i>Ground resolution</i>
<i>Approach 1</i>	17.021.563	0.299 mm/pix
<i>Approach 2</i>	17.138.248	0.291 mm/pix
<i>Approach 3</i>	17.024.734	0.291 mm/pix
<i>Approach 4</i>	17.032.151	0.291 mm/pix

Table 2.21. Final 3D model resolution characteristics

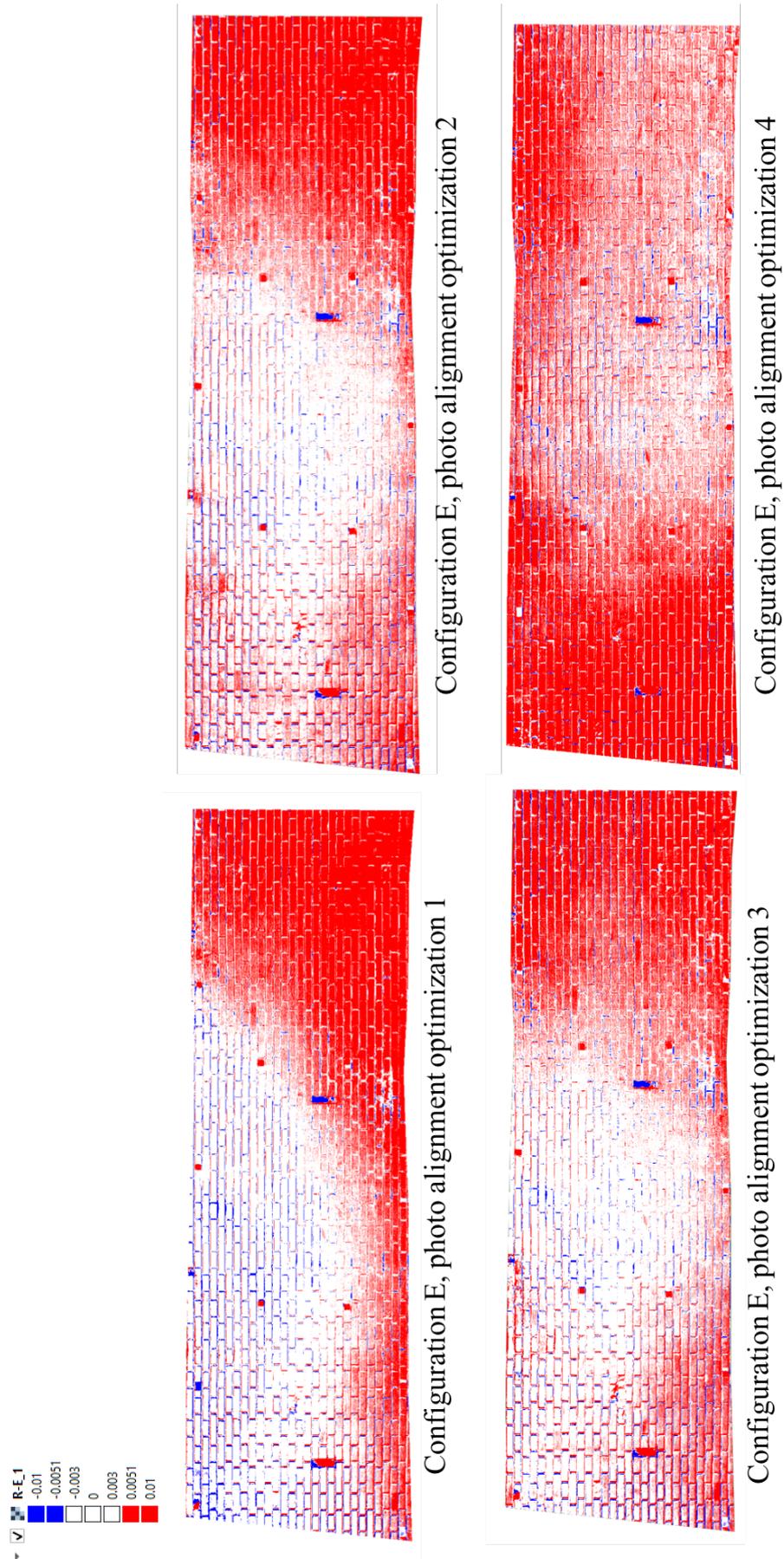


Figure 2.39. Deformation analysis of the models: comparison between DEM from the photogrammetric reconstruction and Riegl DEM.

**Configuration F (Nadiral + Convergent + full facade):**

The introduction of full façade images in the previous camera configuration causes an increase of orientation time from 3 minutes and 54 seconds to 6 minutes and 43 seconds, which, however, remain still a reduced processing time. On the other hand, the reprojection error decreases for all the orientation strategy (Table 2.22) and the Ground Control Point RMSE is subjected to an improvement, except for case 4, that still shows high GPSs errors (Table 2.23). The Figure 2.40 representing the GCPs location errors estimates, suggests that the same behaviour at the model extremities occur, showing Z direction errors of 8 mm. Analysing the model deformation with respect to TLS acquisition, Figure 2.41 shows a clear improvement in the central part, especially in case 1, where the distance from the Riegl DSM is near to zero value for almost the 70% of the surface. Problems at the extremities however remain in cases 2, 3 and 4. The high resolution of the models is confirmed (Table 2.24).

*Camera configuration F – Photo orientation quality*

	<i>RMS reprojection error</i>	<i>Maximum reprojection error</i>	<i>Matching and orientation time</i>
<i>Approach 1</i>	0.845761 pix	31.2483 pix	6 min 43 sec
<i>Approach 2</i>	0.786402 pix	32.9061 pix	6 min 43 sec
<i>Approach 3</i>	0.848329 pix	31.2915 pix	6 min 43 sec
<i>Approach 4</i>	0.848106 pix	31.2915 pix	6 min 43 sec

Table 2.22. Comparison between significant parameters for the photo orientation quality estimation varying the photo orientation approaches.

*Camera configuration F- Ground Control Points RMSE (mm)*

	<i>X</i>	<i>Y</i>	<i>Z</i>	<i>XY</i>	<i>Total</i>
<i>Approach 2</i>	0.280602	0.1765	0.717243	0.331496	0.790144
<i>Approach 3</i>	0.433535	0.165721	0.661822	0.464129	0.808347
<i>Approach 4</i>	3.26931	1.69634	3.45643	3.6832	5.05103

Table 2.23. Ground Control Points RMSE.

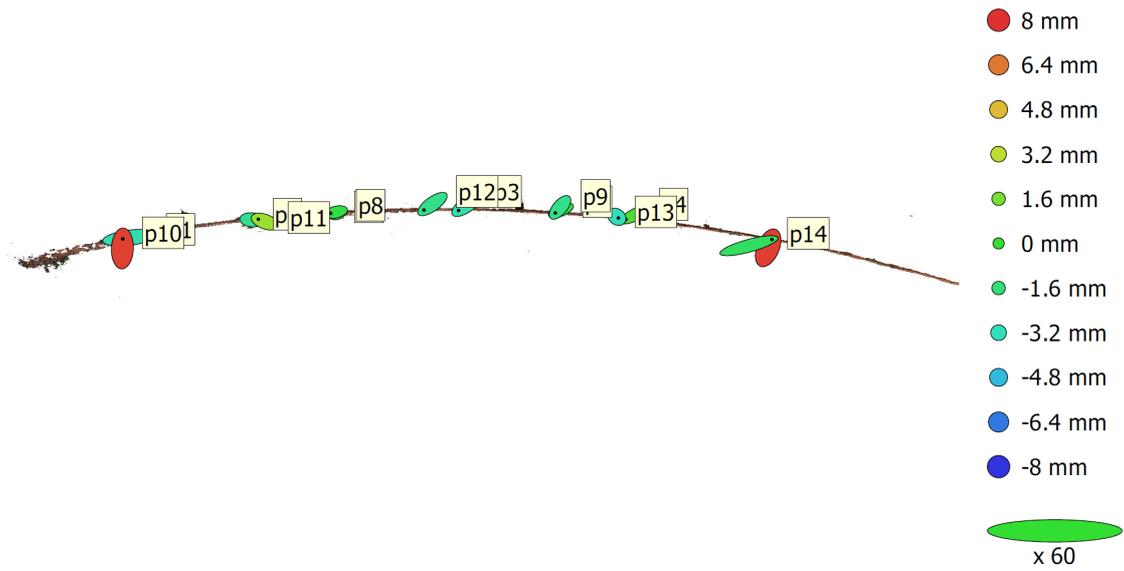


Figure 2.40. GCPs location and estimates (approach 4). The ellipse colour represents the Z error, while the shape represents the X, Y errors.

	<i>Model face count</i>	<i>Ground resolution</i>
<i>Approach 1</i>	16.894.681	0.309 mm/pix
<i>Approach 2</i>	17.022.370	0.291 mm/pix
<i>Approach 3</i>	16.897.393	0.291 mm/pix
<i>Approach 4</i>	16.903.446	0.291 mm/pix

Table 2.24. Final 3D model resolution characteristics

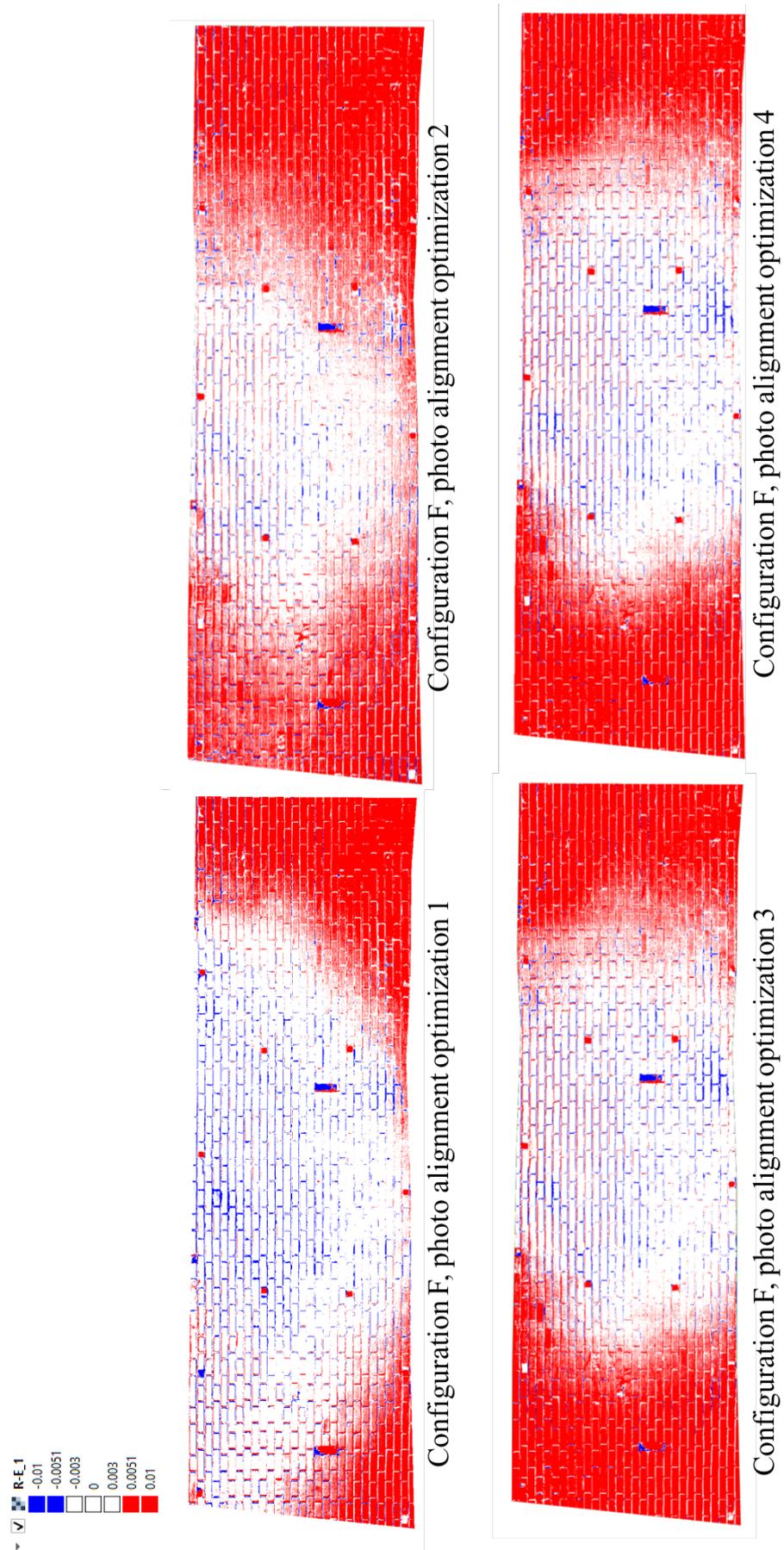


Figure 2.41. Deformation analysis of the models: comparison between DEM from the photogrammetric reconstruction and Riegl DEM.

Considering the photo orientation optimization strategies, the results underline unexpectedly the scarce effectiveness of numerous GCPs (photo alignment 4) in eliminating the deformation at the extremities of the model. In configurations A, B, C, D and E this effect is clearly visible from the Figure 2.42. The outcomes of photo orientation algorithm with approach 4, differs considerably from the expected results that should provide a more stable geometry reconstruction due to the GCPs location: this inconsistency, which have badly influenced the photo orientation optimization, can be attributable to the high degree of constraints assigned to the process (14 GCPs), that have resulted in overabundance of parameters. An exception can be made for configuration F: the area without significant deformations (white central area) is the largest area obtained from this experimentation, meaning that the employment of nadiral, convergent and full façade images causes in general the reconstruction of more reliable 3D models. In all the camera configurations, the model generated with the photo orientation approach 1, that simulates a quick survey, gives in general good results comparable with those obtained with approaches 2 and 3 that use GCPs.

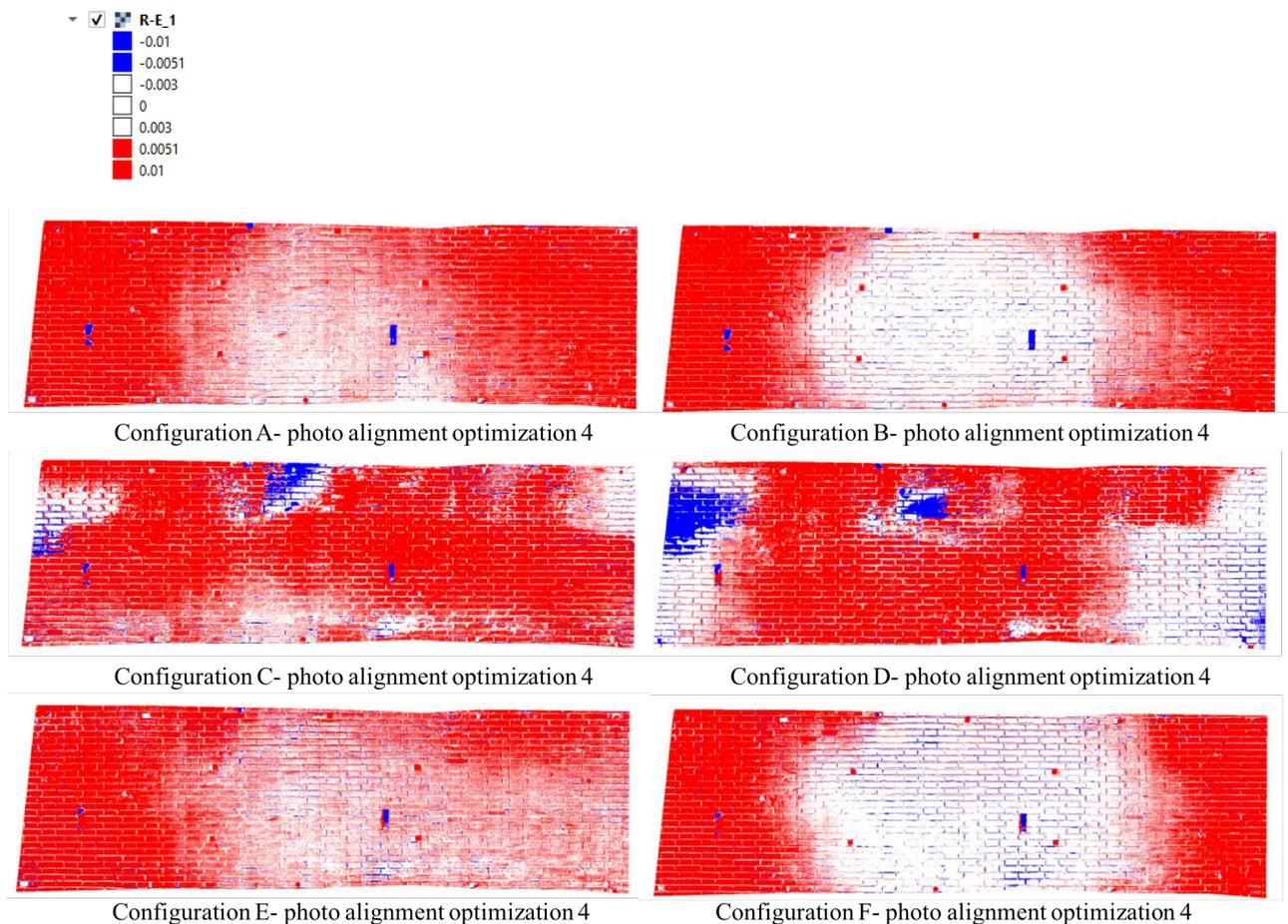


Figure 2.42. Deformation analysis of the models considering photo orientation approach 4 varying the camera configuration: comparison between DEM from the photogrammetric reconstruction and Riegl DEM.

Finally, considering the camera configuration issue which is the central issue of this section, it is reasonable to affirm that a photogrammetric survey that involves only convergent images does not provide a reliable and accurate reconstruction of the scene without anomalies and deformations. For this reason, it is convenient to join nadiral uniformly distributed images in order to stabilize the

process of reconstruction. Full façade images, in presence of weak camera configuration geometry (only nadiral or only convergent images) play a minor role, indeed worsening the reliability of the reconstruction; however, if used in a complete and rigorous acquisition process, they become fundamental to ensure low deformed models.

### **3. RANGE BASED METHODS: FROM LASER TO STRUCTURED-LIGHT SCANNER TECHNOLOGY**

#### **3.1 History and developments**

Optical range-based methods developments over years have always be influenced by the progresses of the electro-optical technology and the disposal of new electronic components and devices (Besl, 1988). First researches in the field of optical contactless sensing technologies dated back to the end of '70s, when considerable American and Canadian scientific laboratories began to study the light projection technology as functioning principle for 3D sensing devices. The aim was to overcome the limits of contact probes and exploiting the use of light beam for measuring systems in mechanical manufacturing industry. In early '80s, the National Research Council (NRC) of Canada began to deepen the studies related to 3D surveying, experimenting new active sensors based on triangulation principle: the first prototype of modern triangulation laser took advantage of the advent of the digital technologies that led to a breakthrough with respect to old analogic devices, such as Lateral Effect Photodiode systems (LEPs), that despite having good acquisition performances were difficult to calibrate and scarcely accurate. The first NRC triangulation-based sensor exploited the introduction of the CCD arrays and used a linear 2048-element sensor. This digital solution ensured higher accuracy and stability to the range data: coupled with real time digital peak detection algorithms, led to range resolution 10, and successively, 100 times higher than the analogue one (Blais, 2004).

The '90s were the years devoted to the production of a wide choice of devices thanks to the availability of relatively low-cost components and techniques. The instrumentation of those years reached high measurement resolution and ranges; jointly, software-level researches started, aiming to develop powerful programs to manage the data acquired with the new instrumentation. As a result, the interest in 3D contactless active techniques increased: the field of mechanical industry, that traditionally has made use of contact probes for shape acquisition, has now become interested in optical sensors to be coupled with the traditional instrumentation. An example are the modern CCMs, that mount both optical 3D measurement sensors and contact probes. Beyond the manufacturing industry however, range-based sensors have recently found wide application on several uncommon fields, from Cultural Heritage to medicine, from civil engineering to forensic. Related to this employment evolution, another kind of development has occurred, that regards the possibility to produce low-cost devices with increased portability and ease-of-use, that are safe for humans and for the materials to be surveyed (Sansoni et al., 2009). These requirements regarded mainly the modern Terrestrial Laser Scanners and the structured-light scanners, whose characteristics and functioning principles will be described in next paragraphs.

#### **3.2 Measurement principles**

Optical active sensors take advantage of light emission for the 3D shape measurement, which can be coherent, such as laser beams, or incoherent, and this is the case of structured-light. Depending on the number, the shape and the codification of light spots emitted (single point, lines, fringes), spatial resolution, speed and accuracy may substantially vary, such as the application field and the scale of the scene to be acquired. Measuring techniques of active sensors mainly vary depending on the sensing characteristics mentioned above, that influence their functioning principle. There are two

main typologies of range-based techniques: the one based on the triangulation principle and the one based on the distance measurement. Next paragraph will deepen the triangulation systems.

Independently from the measurement principle however, the first rough data of the range-based sensors acquisition is often a point cloud, for which the post-processing strategies introduced in the paragraph 2.5.3 are still valid and widely adopted.

### 3.2.1 Triangulation systems

Classic triangulation principle is at the basis of this technology: it states on simple geometric rules, and in the past, it was used by Greeks for basic geodesic measurements in the field of navigation and astronomy. The process aims to define the location of a point P through the knowledge of the angles at the base of the triangle determined by the intersection of straight lines passing through the point and having a known base B (Figure 3.1).

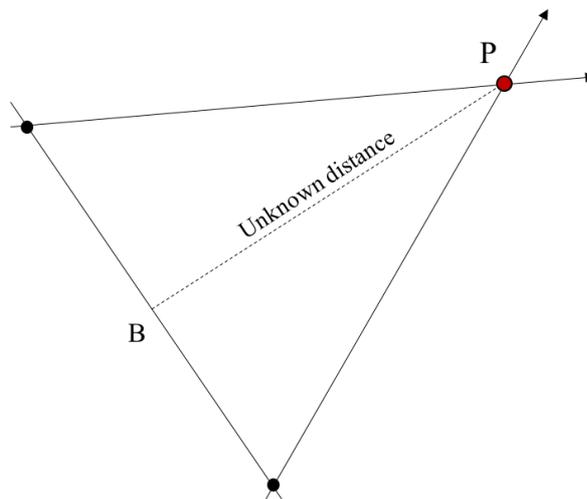


Figure 3.1. Triangulation principle.

The first type of light that has permitted the realization of a 3D triangulation device is the LASER, a coherent emitted light. LASER is the acronym of Light Amplification by Stimulated Emission of Radiation, whose operating principle is the transformation of chemical and electrical energy in a high-energy radiation. Starting from the natural light characteristics and behaviour, it is well known that it consists in photons, which are spontaneously emitted by atoms: the latter one are composed by electrons and protons orbiting around the nucleus and occupying precise energy positions which can change only if the atom is stimulated by energy. When a certain amount of energy is added, the electrons jump from their position to another one farer from the nucleus and proportionally to the energy applied. Instead, when an electron moves closer the nucleus, the atom emits a photon, which is energy in the form of electromagnetic quantum.

The energy carried by a photon is proportional to its wavelength. The process that allows to obtain stimulated photons emission is called *population inversion*: it is a procedure that consists in stimulating the atoms such that the electrons occupy a position characterized by an energy level higher than the minimum allowed. In this condition, the electrons tend to return to the lowest energy level generating energy (light) without any external intervention. Considering the light used for the atoms stimulation as a beam of photons, when one of them hits the atom, a certain amount of energy is emitted in the form of light with the same wavelength and phase of the photon. This results in the

generation of a new photon, identical to the previous one which, in turn, will hit another atom generating another photon, and so on. This cascade of stimulated emissions gives the LASER.

Figure 3.2 shows an example of device for the LASER generation.

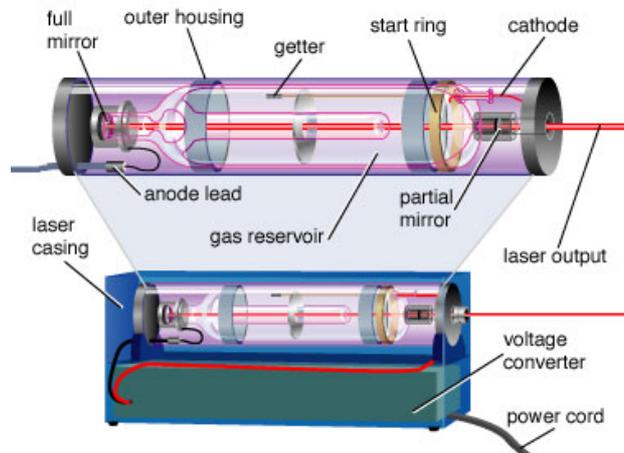


Figure 3.2. LASER emission device (Encyclopaedia Britannica, Inc)

Inside the gas reservoir, there can be two or more elements, like helium and neon, which are excited releasing photons. The device length influences the light intensity, the cascade effect in fact has an intensity that become higher and higher with the increasing of the distance that light must cover; that is the reason why the device develops especially in the longitudinal dimension. The mirrors have the function to increase the distance to be travelled by light, thus to increase the beam intensity, making the light bouncing between them before leaving the device. The output emitted through the LASER hole is a thin laser beam whose radius depends on the distance travelled by light.

Therefore, the constitutive element of a LASER is the electromagnetic energy, the same as the natural light, with the difference that, in the latter case, the photons propagate in all the directions and with different wave length, while in the case of LASER, the result is a monochromatic light beam characterized by only one wavelength. It is also *spatially coherent*, namely all the electromagnetic energy photons emitted are in phase; this characteristic allows the LASER to stay in focus when projected on a scene but, on the other hand, it implicates some speckle when in presence of rough surfaces, that is source of noise. Among the advantages provided by LASER, there is a very small beam divergence due to the same propagation direction of photons. In addition, the light produced by LASER is highly bright. All these features make the optical systems able to transport the beam and therefore to concentrate the energy that it transports.

The triangulation devices, in general are constituted by a light source that emits a LASER beam and a solid-state image sensor (CCD or CMOS) as a receiver. They are included in the *scanner head* and are placed at a fixed distance, baseline  $b$ . The emitted LASER beam, deflected by a mirror, impinges on the object surface at point P, whose image is acquired by a camera composed by lens and sensor. As shown in Figure 3.3, the light source, the projection centre on the sensor and the point P are the vertices of a triangle of base  $b$ . This triangle follows the rules of the triangulation: the knowledge of the angles  $\alpha$  and  $\Delta\beta$  are necessary for the determination of the coordinates of P.  $\alpha$  angle is directly measured, while  $\Delta\beta$  is measured through the theorem of opposite angles, assuming  $\alpha$  constant, for an incremental change of distance  $\Delta Z$ .

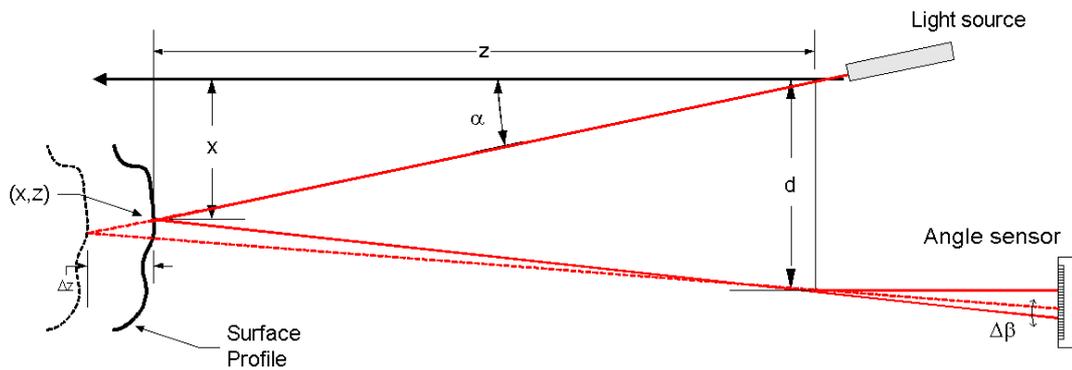


Figure 3.3. Single point triangulation-based laser scanner functioning (Beraldin et al., 2000)

This is the configuration of a *single spot laser scanner*. Setting a local coordinate system where:

- $x$  axis contains the emission point of the laser and the point on the sensor hit by the emitted light
- $z$  axis lies on the plane containing the point P and the  $x$

through simple trigonometric relationships, the range  $Z_P$ , that is function of  $\alpha$ ,  $\beta$ , and  $b$  is calculated by the following:

$$Z_P = \frac{b}{\tan \alpha + \tan \beta}$$

Single spot scanners measure the surface performing a scanning of the scene; however, some limitation exist, due to the relationship between the Field of View (FoV), resolution and uncertainty and the shadow effect. During '80s, in the NRC laboratories, a crucial method to overcome the limitations, called *synchronized scanning*, has been proposed and then presented by Rioux in 1994. This principle uses fast scanning mirrors: the idea is to synchronize the sensors with the laser source, namely the laser projection and detection, to achieve larger Field of View with a small triangulation angles maintaining high precision (especially in  $z$  component) and eliminating the shadow effect. Today, this solution results in a considerable reduction of the scanner head and optimization of the CCD sensor, whose length can be used totally. However, the cost is usually high due to the onerous price of specialized mechanical components.

The configuration explained above, can be extended from a single point to a laser line, i.e. a set of aligned points that constitutes a segment. The device, called *slit scanner* (Figure 3.4), is the most widely used triangulation-based 3D scanner since allows the detection of a complete profile of points in a single scan. The information related to the surface geometry, are inferred from the analysis of the profile deformation, that is direct function of range. Slit scanners are less expensive than single spot scanners due to the introduction of smaller and cheaper CCD and CMOS sensors. On the other hands these devices have some limits regarding the Field of View, that is kept usually low (20-30 deg) to improve the depth resolution, as well as the influence of the ambient light on the system. Even

though the introduction of narrow-band optical interference filters, the application of slit scanners outdoor has still some criticalities (Beraldin et al., 2000; Blais, 2004; Guidi et al., 2010).

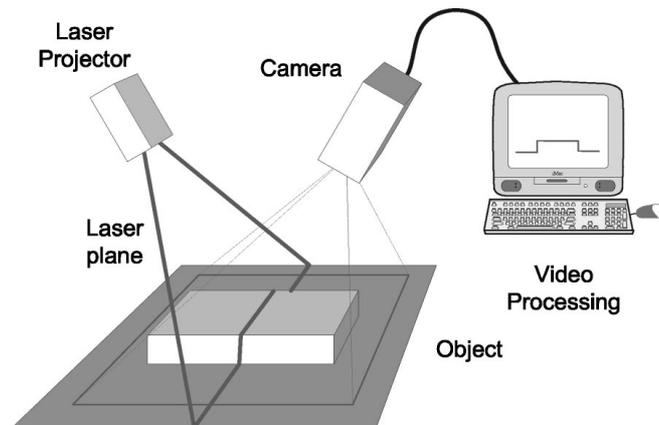


Figure 3.4. Slit scanner scheme (Blais, 2004).

In general, triangulation systems are designed to be applied with a low operating range, due to the physical limitations on the base length. Based on the operating range, a distinction between two types of sensors can be made: short-range sensors can cover a distance less than 0.5 m, while medium-range from 0.5 to 2m. Among the triangulation systems, *pattern projection* devices deserve special attention: they can be considered slit triangulation scanners where slits, instead of moving horizontally until the complete scanning of the surface, are projected all at the same time constituting a codified *pattern*. The peculiarity of these system, beyond the pattern projection, is the demodulation of the signal. Next paragraphs deepen this technology, focusing also on the calibration issue of these systems.

### 3.2.2 Structured-light systems (SLS)

Structured-light systems are sensing devices that exploit the triangulation principle for the acquisition of 3D information about a scene. Conversely to single spot or slit scanners, they do not make use of laser light. Instead, a SLS project simultaneously one or more bi-dimensional *pattern*, known as *code*, on the object using a visible non-coherent light. The device is generally composed by a projector and a camera: the projector projects the coded pattern on the scene and the camera captures a series of images of it. Compared to a stereoscopic system, the projector can represent the second camera placed at a fixed distance (baseline) from the other one (Figure 3.5).

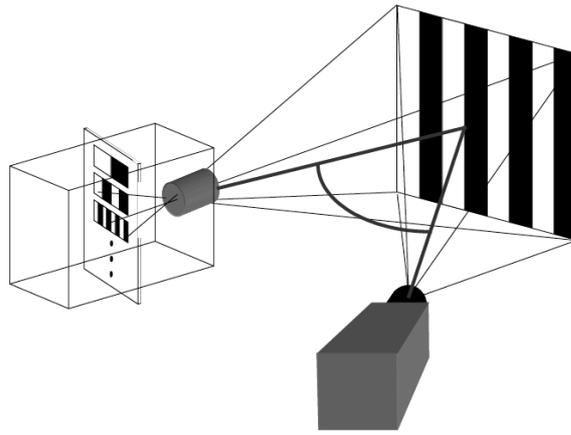


Figure 3.5. SLS functioning principle (Acka et al., 2006).

The model proposed above is the simplest configuration of a SLS, where a single camera is coupled with a single projecting device. More recent widespread devices however, take advantage of two cameras of the same focal distance to provide over-determined mathematical triangulation model. One great advantage of structured-light technology is therefore the ability to measure million points within a single view measurement; in addition, thanks to the non-coherent light source type, any speckle effect is avoided. On the other hand, the range of a SLS is quite low, since light intensity tends to decrease rapidly with distance from the source (Jecic and Drvar, 2003).

Two codification models are widely used for SLS depending on the type of the projection and the codification, the *grey-code* method and Moiré technique (Valkenburg and McIvor, 1998; Salvi et al., 2003; Chen et al., 2009; Guidi et al., 2010;).

### **Gray-code method**

The technique projects multiple binary-coded images (Gray-Code) which progressively increment the resolution. Considering a sequence of black and white stripes, the method describes uniquely  $2^n$  different direction of projection, where  $n$  is the number of projected images characterized by a unique pattern (the code length). The first projected pattern is a single alternation of black and white strips; as the method proceeds, the following patterns projected are characterized by a progressive doubling of strips until the maximum frequency is reached (Figure 3.6 and 3.7). Each pattern projection lasts few tenths of seconds: the camera acquires, registers and saves each correspondent image of the pattern while the procedures runs.

Usually, the tighten steps vary from 7 to 10 (however it depends on the camera resolution). The adjacent strips are codified through a binary code and differ for one bit (0/1), where 0 stands for black and 1 for white. Therefore, each pixel is characterized by a sequence of 0 and 1 depending on its grey level (a sequence of 0 indicates a black zone, while a sequence of 1 a white zone). If the SLS works with 10 levels, the sequence is composed by 10 digits.

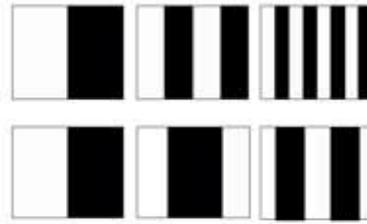


Figure 3.6. Tightening of the pattern stripes (Salvi et al., 2004).

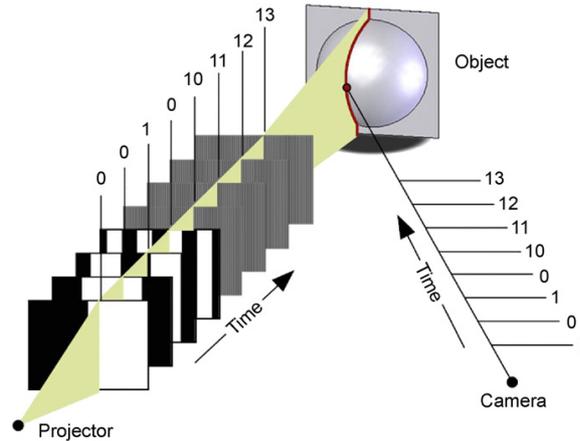


Figure 3.7. Encoding scheme in function of time (Chen et al., 2009).

### *Moiré technique*

The Moiré principle is based on the measurement of the signal frequency difference and specifically, it estimates the periodicity of low frequency variation of two amplitude modulated signals interacting together. The technique consists in analyzing the interferences generated by the superimposition of two similar patterns (Figure 3.8): Moiré fringes are obtained by illuminating the object with a light passing through a grating and acquiring the scene with a camera through a matched second grating. The patterns modulate the amplitude of the projected light, as well as the camera demodulates the superimposed fringes. The interference fringes are strictly related to the range, namely they are proportional to the surface development in depth. In practice, the grey level depends on the phase shift: the higher the shift, the whiter the grey level. Other methods use software to demodulates the signal or stereoscopic systems for the projected pattern detection.

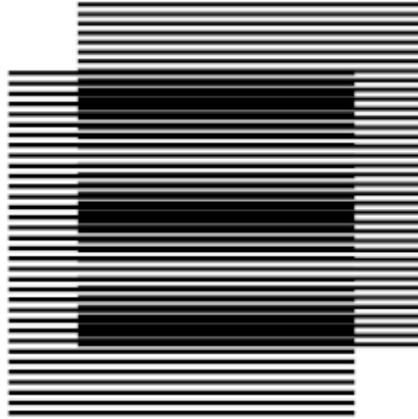


Figure 3.8. Moiré effect: two slightly different pattern generates a third pattern that represent the interference fringes (<https://bowiestie.files.wordpress.com/2010/02/image024-11.gif>).

In general, the advantages of structured-light 3D scanners are the acquisition speed and the great versatility: some recently developed SLS are capable of scanning moving objects in real-time and, for their portability and accuracy can be employed in several different fields. Examples of application fields are: archaeology, civil engineering, medicine, industry. In general, structured light scanners have proven to be a fundamental tool in the reverse engineering process thanks to the possibility to eliminate the need of a scratch to design a product and to provide a fast and accurate path for its development.

### 3.2.3 Terrestrial Laser Scanners

Terrestrial Laser Scanners (TLS) are long range sensors working with a range greater than 10 meters, used for the acquisition of large scenes. These systems derive from Total Stations, optical instruments for the measurement of angles between designated visible points in the horizontal and vertical planes, integrated with electronic distance measurement. If Total Stations require an operator to individuate points of interests in the space, TLS have the capability to automatically orient the laser pulse on a predefined range of horizontal and vertical angles, thus selecting a specific area in front of the instrument. Digital encoders then return the angular estimations and the receiver estimates the range. Most of modern TLS rely on angular movements that can reach  $360^\circ$  horizontally and almost  $180^\circ$  vertically, thus allowing the acquisition of a volume of spherical shape from a fixed single station. In general, the range systems based on distance measurement are known as LiDAR (Light Detection And Ranging) whose functional scheme involves a light source and a receiver that compose the scanning unity (Figure 3.9). However, in topographic field the acronym LiDAR refers to a specific type of airborne laser scanner. Here the ground based LiDAR is considered, known indeed as TLS.

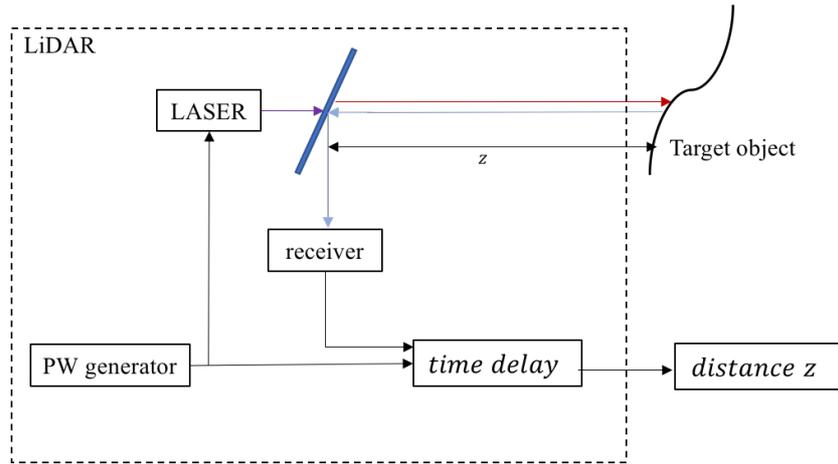


Figure 3.9. Functional scheme of LiDAR

The primary classification widely adopted in literature distinguishes between instruments that employ phase shift measuring and those that rely on the time of flight (ToF) estimation (Shan and Toth, 2017). The phase shift laser scanners are mainly involved in short-range measurements (from 50 to 150 m): the signal emitted by the instrument is a continuous laser beam whose amplitude is modulated assuming a sinusoidal wave pattern. The measuring principle rely on the comparison between the transmitted and the received versions of the transmitted beams and evaluating the phase difference between them. The actual comparison take place at the laser rangefinder (Figure 3.10). The carrier signal, in turn, is modulated through the superimposition of a measuring wave pattern on it, so that the short wavelength ( $\lambda \approx 1 \mu m$ ) can be easily measured. Knowing the phase difference (or the phase angle  $\phi$ ) and the fractional part of the total distance  $\Delta\lambda = \left(\frac{\phi}{2\pi}\right) \cdot \lambda$ , the final distance  $R$  is given by the following:

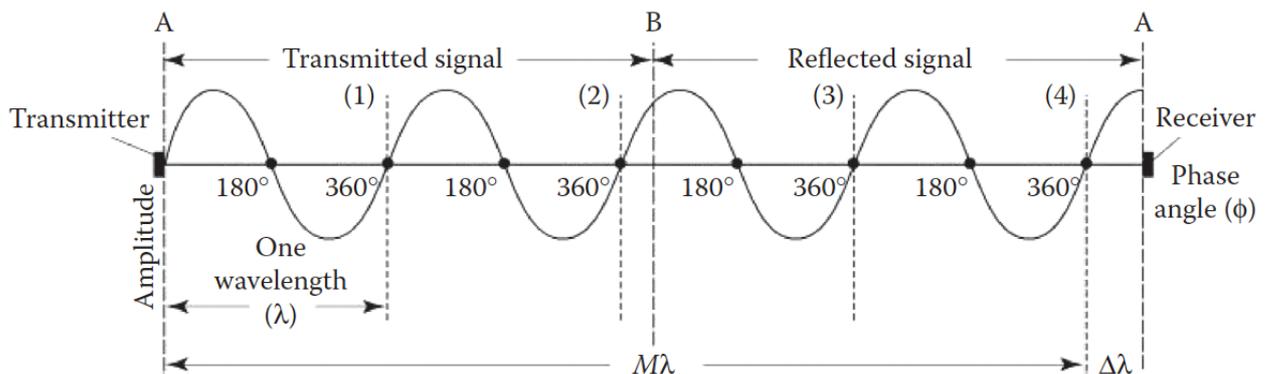
$$R = \frac{(M\lambda + \Delta\lambda)}{2}$$

where:

$M$  is the integer number of wavelengths;

$\lambda$  is the known value of the wavelength;

$\Delta\lambda$  is the fractional part of the wavelength.



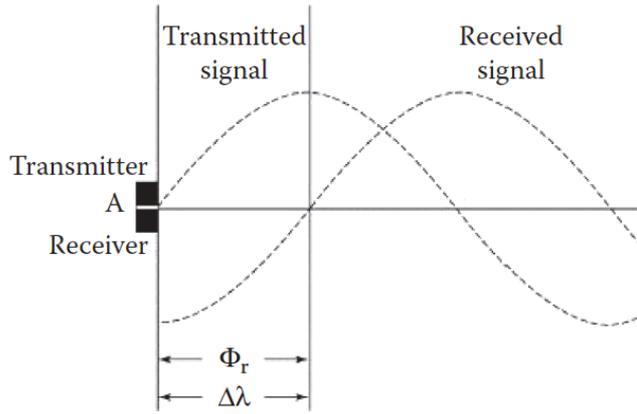


Figure 3.10. Continuous sinusoidal transmitted wave (top) and phase comparison between two signals at the range finder in A (Shan and Toth, 2017).

ToF laser scanners instead, take advantage of an intense pulse of laser radiation, and for this reason they also known as Pulse Wave (PW) systems. The measurement technology is based on the time delay of a coherent laser pulse emitted from the instrument, and reflected back by a surface to the receiving detector. Through the round trip  $\tau$ , the distance of the scene point from the instrument center is calculated as follow:

$$R = c \frac{\tau}{2}$$

where  $c$  is the light speed, that can be assumed equal to  $3 \cdot 10^8 \text{ m/s}$ . The measurement of the distance  $R$  jointed with the knowledge of the rotation angles of the scanner, allows the estimation of the 3D coordinates of the object points. The positioning operation is fully automated thanks to a motor that changes automatically the beam direction following planned steps. In this work, ToF systems will be analyzed.

ToF devices generally consist in:

- *Semi-conductor laser emitter*, that emits infrared signal pulses directed and conducted by a lens;
- *Rotating mirrors system*, that deflects the laser beam;
- *Receiving lens*, that intercept the part of the laser beam reflected by the object surface;
- *Receiving photodiode*, that emits an electric signal representing the laser pulse reception. In order to facilitate the collection of 3D data, recent studies suggest the employment of array of *avalanche photodiodes*, which allows a more rapid collection of 3D data.
- *A quartz stabilized watch*, that measures the time delay of the signal;

The functional scheme of a ToF system is represented in Figure 3.11

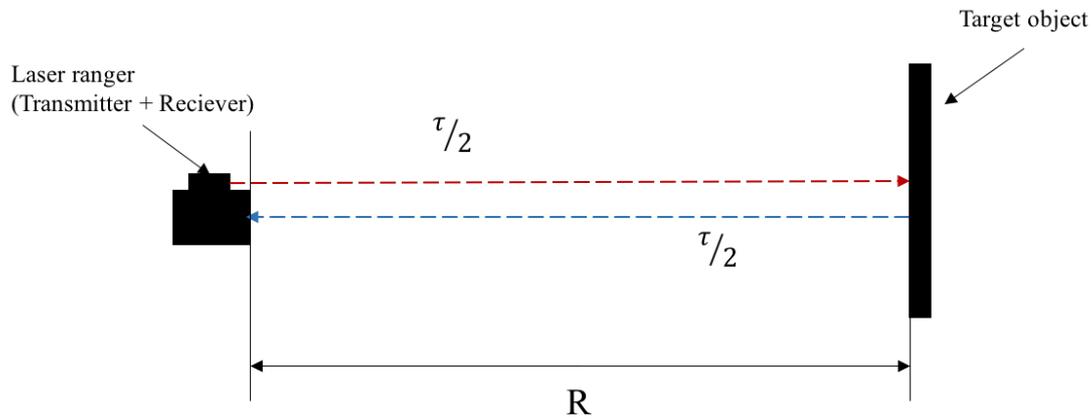


Figure 3.11. Functional scheme of a ToF system: the estimation distance  $R$  is performed through the knowledge of the round trip  $\tau = (2 \cdot \tau/2)$ .

The ToF systems tend to be more stable than the triangulation-based sensors. The components stability, together with high-frequency bandwidth electronics and constant group delay, allow the ToF long range sensors to reach very high resolution values maintaining short pulses (notice that 1 mm resolution would imply a time delay of about 3.33 picosecond). However, in order to handle large bandwidths, sensitivity to signal strength variation need to be minimized, and this is possible utilizing complex electronic components. As a result, the cost of these systems increases, making the ToF systems more expensive than triangulation-based sensors (Beraldin et al., 2000). ToF systems are less accurate than SLS (the accuracy is in the range of some millimeters to few centimeters) but, on the other side, they offer the great advantage to work in a wider distance range (the operative distance of ToF systems varies from 10 meters up to few kilometers) adopting high acquisition speed (up to 2 million points per second). However, the range accuracy is relatively constant for the whole volume of measurement (Blais, 2004).

The accuracy of laser scanners is an articulated issue: in general, the values reported on the datasheet provided with the instrument are scarcely reliable, since they refer to laboratory tests in ideal conditions. Observations based on experience show that the accuracy, indeed, varies from instrument to instrument and depends on the individual calibration experience and the attention paid handling the scanner itself. Another important consideration when treating laser data, regards the term “accuracy”. Consider a series of point clouds of one scene: they are sets of considerable numbers of points acquired singularly by the sensor that very improbably will be the same in each cloud. Then, it is reasonable to speak about *repeatability* of the measure and is important to be aware that the laser measures the *redundancy* of the entire surface (not the single point). For this reason, each point has a single value of accuracy, which is determined by the error propagation related to the factors influencing the single measurement. (Boehler et al., 2003) show and test a set of fundamental parameters for the accuracy evaluation of a laser scanner device.

The angular accuracy, influencing directly the 3D point coordinates computation; the range accuracy, affected by systematic constant error component and noise (accidental error); the resolution, to which contribute the increment of angle between two successive points and size of laser spot on the object (Lichti and Jamtsho, 2006); the edge effects that cause wrong points in the vicinity of edges; and the ability of surface to reflect light (albedo).

***Range accuracy: a case study***

Considering the range accuracy, several works has been published, which investigate the distance measurement error sources (Soudarissanane et al., 2011; Tan et al., 2018;). The improvement of TLS measurement accuracy in fact is of fundamental importance, since it determines the accuracy and reliability of output data (point clouds) and the good performances of the instrument. The achievable distance precision of a ToF system however, depends on various phenomena: instrument mechanics, environmental conditions such as temperature and atmosphere condition (dust, steam), scanning geometry and configuration (distance and incidence angle), and the object surface characteristics. In general, noise effects occur when considering range measurement, and they are caused by optical phenomena related to the reflection of a laser light. They are related to surface material, surface roughness, incidence angle, energy distribution within the footprint, and surface reflectance; even if the surface texture implies errors in the order of millimetres, the properties of the surface material affects more significantly the distance deviations, that can assume errors values in the order of centimetres and decimetres (Bohler et al., 2003; Dorninger et al., 2008). Fort this reason, several interesting works have deeply investigated the effect of reflecting surfaces properties on ToF laser scanner measurements, as they badly affect instrument calibration, relative registration of points clouds, surface modelling precision, and georeferentiation of models (Lichti and Harvey, 2002; Zámečníková et al., 2014;).

In these works, the reflectorless nature of the rangefinder is considered as the cause of possible range errors, that arise due to pulse attenuation by the reflecting surface. The influence of the surface material properties on the laser signal can be treated considering two aspects of the issue. The first one regards the maximum achievable range when scanning high reflective materials, which is considerably higher than low reflectivity materials one (Figure 3.13); and the second is the range delay error  $\delta\rho$  caused by the amplitude attenuation of the reflected signal in presence of reflective materials. Moreover, considering reflective surfaces, it is important to distinguish between two reflection components, the diffusive and the specular; stated that specular and diffuse reflection exist in numerous natural surfaces, the specular component is dominant in particularly smooth and reflective materials such as water, ice, leaves, metal, marble, and plastic. Here, retro-reflective targets are considered and in particular, some issues connected with them are presented.

The case study refers to the survey of the brick-wall presented in Chapter 2 (Figure 3.12). For the acquisition, Riegl VZ400 is used (Figure 3.13; Table 3.4 for specifications). The survey geometry involves one centred scan position, at a distance from the wall of 5 m. The point cloud resolution obtained is 1,5mm (Figure 3.17).



Figure 3.12. The brick wall acquired by Riegl VZ400. Dimensions: 8x2,5m.



Figure 3.13. Riegl VZ400 mounted on a tripod (<http://www.riegl.com>).

<b>Minimum Range *</b>	1.5 m
<b>Maximum Range *</b>	>600 m
<b>Accuracy</b>	5 mm
<b>Repeatability</b>	3 mm
<b>Max Measurement Rate (meas./sec)</b>	122 000
<b>Laser Wavelength</b>	Near Infrared
<b>Laser Beam Divergence</b>	0.3 mrad
<b>Angular Step Width</b>	0.0024° ≤ Δθ ≤ 0.288° (vertical) 0.0024° ≤ Δφ ≤ 0.5° (horizontal)

Table 3.4. Laser Scanner VZ400 specifications. \* the measurement range values are dependent on Laser Mode (Long Range/High Speed) and target reflectance. See Figure 3.23

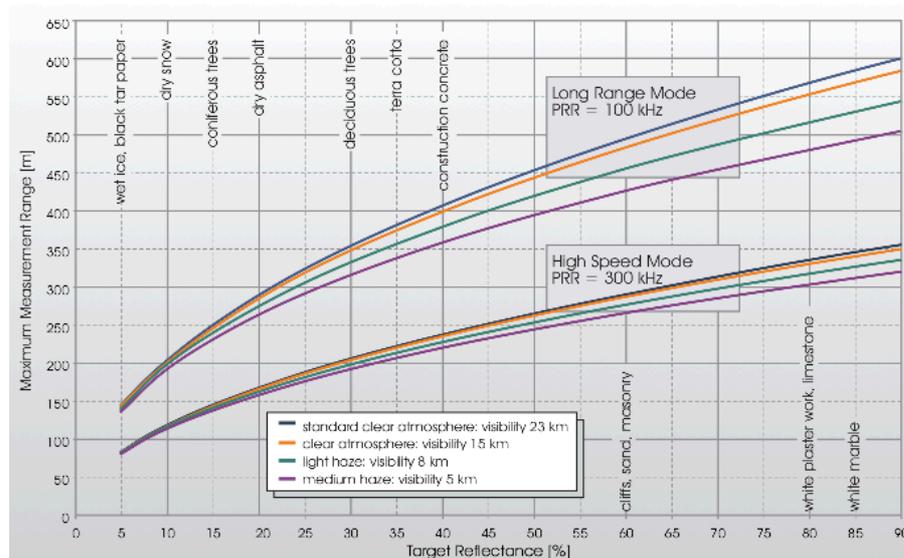


Figure 3.14. Target Reflectance/Maximum Measurement Range Chart (two separated trends for the Laser Modes) (<http://www.riegl.com>).

The target used are those shown in Figure 3.15: they are small retro-reflective targets characterized by a white square of reflective material with a black marked cross in its centre. 12 retro-reflective targets have been attached on the surface of the test wall, as the scheme of Figure 3.16 shows.



Figure 3.15. Example of retro-reflective target used during the surveying with Riegl VZ400.

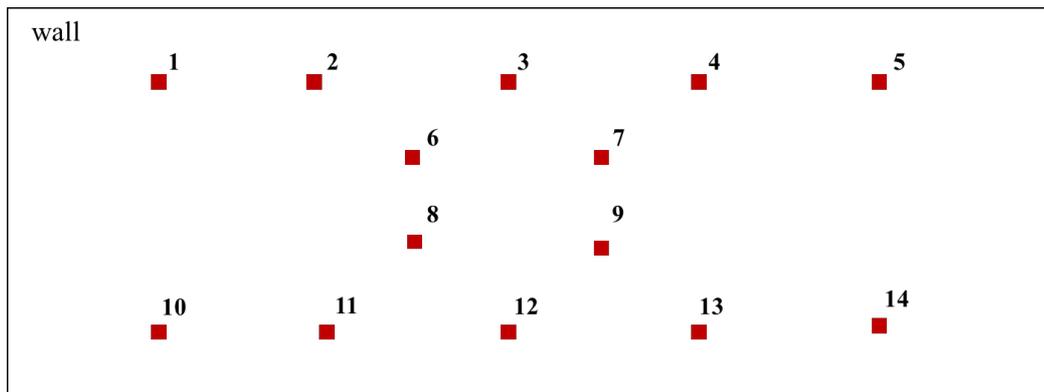


Figure 3.16. In red, the retro-reflective targets.



Figure 3.17. Point cloud of the wall, 1,5 mm resolution.

As expected, the processing of scanning data has revealed irregularities in correspondence of the targets. Specifically, as Figure 3.18 shows, the white points (representing the targets) are positioned outside the wall plane, along the direction connecting the instrument and the target; this means that the surveyed targets are nearer the scanner, reason why this effect is called *shortening range effect*.



Figure 3.18. Shortening range effect in correspondence to a retro-reflective target.

The effect can be explained considering both the saturation effect and the specular reflection component. The saturation of the signal into the internal sensor (photodetectors), in fact, can lead to a truncation of the returned pulse due to extremely high amplitude of the returned signal. The specular reflections, on the other side, can affect significantly the amplitude, the width and the returned signal shape, especially if incident angles are close to zero; this leads to a modification of the returned waveform, that is generally characterized by a peaked shape flatter than the emitted one. This deformation affects significantly the capability of the internal sensor to correctly determine the impulse arrival time, resulting in distance measurement error (Pesci and Teza, 2008). Further experiments at different operative distances involving targets of different reflective material are necessary in order to investigate in deep this phenomenon.

### 3.3 Active sensors calibration issues

With the term calibration, we refer to the procedure aiming to the extraction of the internal orientation parameters of the instrument in order to check and adjust the measurements acquired. This is possible if errors, introduced in data acquisition phase, are known and if the behavior of the instrument is predictable and verifiable. Therefore, calibration allows the user to monitor the instrument performances along time and eventually change its settings with respect to external conditions. The calibration procedure must be performed before each measurement stage: in fact, it influences the parameters needed for the 3D modelling thus the overall accuracy of the measurements. The process, however, is feasible only if the parameters behavior can be described through a mathematical model.

When dealing with active sensors, the calibration procedure becomes an issue, since the calibration algorithms, based on certain mathematical models, are not provided to the user, who cannot perform any calibration procedure, except using specialist instrumentation. Calibration issue has been faced in several works over the years and all of them have treated common problems, such as: the lack in specific international standards, through which verify the specifications declared on the scanners datasheet; the adoption, from producers, of internal reserved standards and procedures; and the absence of a standard calibration method. In general, the aim is to find a methodology which could ensure an accurate estimation of the measured points coordinates in a definite, coherent and repeatable way. Thus, the problem is not the calibration of the instrument itself, but its

*characterization*, namely the definition of the operating parameters, in such a way that its performances can be estimated in different working conditions. The aim of the *characterization* is to determine the quality of the acquired datum, and the parameters which allow to study it are basically *resolution*, *accuracy* and *precision*. These three parameters can be represented as follows (Figure 3.19)

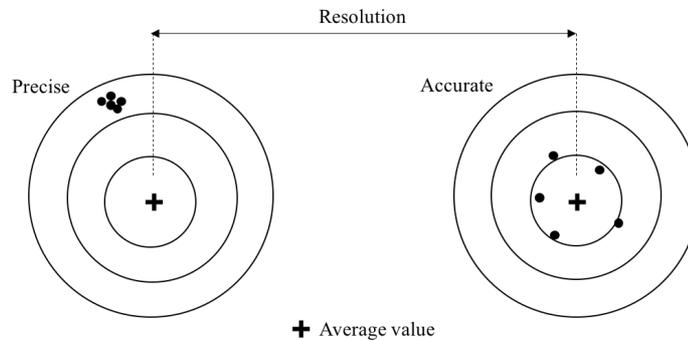


Figure 3.19. Graphical representation of precision, accuracy and resolution: the two circles represent two separate sets of measures.

They are defined in the Accuracy Standards for Digital Geospatial Data ASPRS as follows.

- Resolution: the smallest observable variation of the object measured. In metrological terms, and when considering instrumentation for survey activity, is the smallest spatial step sample available. Considering the points belonging to a 3D space, it is convenient to distinguish between:
  - Horizontal resolution, in x and y direction (considering horizontal the plane orthogonal to the axis of the instrument): x-resolution could be different from y-resolution if the scanner moves.
  - z-resolution
- Accuracy: the error between the measure and its real value
- Precision:  $\sigma$  of the set of measures with respect to the mean value

Resolution, accuracy and precision are the most important parameters of this kind of instruments because they strongly contribute to the good success of the 3D model reconstruction. Usually the producer provides a datasheet coupled with the instrument, that lists these metrological characteristics. The datasheet, however, in most of the cases, is incomplete and ambiguous and generates lack of information regarding the origin and the estimation method of these data. Furthermore, a good characterization depends also on the goodness of the methodology adopted on the laboratory where to perform the experiments, which must ensure a proper temperature, a good illumination and the absence of any external interferences; and also on the objects chosen for the test: they must be certified, namely their dimension and geometry must be known through CMM (Coordinate Measurement Machine) or from the object producer.

The methodologies proposed in recent works (Guidi et al. 2010) and (Genta et al., 2016) provide for the testing of different objects characterized by defined geometry and measured

through CMM, such as a set of steps, a set of solids, a parallelepiped block, a reference plane and a ball plate (Figure 3.20).

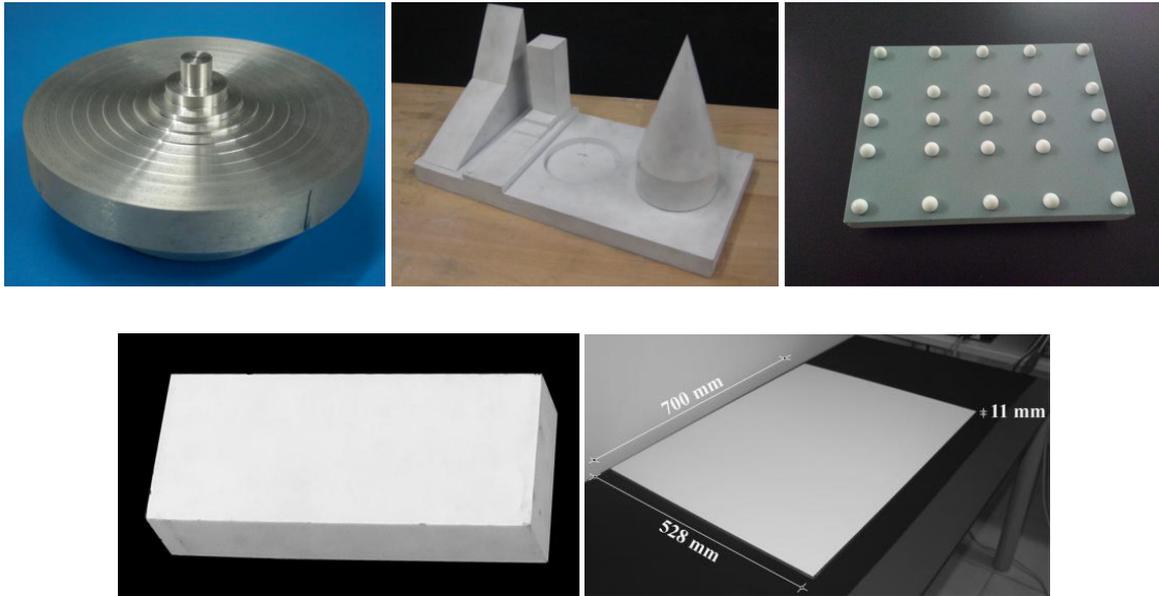


Figure 3.20. Testing objects. First row, from left to right: a set of steps, a set of solids and a ball plate. Second row: a parallelepiped block and a reference plane (Guidi et al., 2010; Genta et al., 2016).

### 3.4 Calibration of structured-light scanners

SLS structure consists basically of a light projector-camera pair that fulfills the same function of a stereo system, where relative orientation is supposed to remain the same. Therefore, the calibration of these systems is a fundamental preliminary step that leads to a correct 3D reconstruction of the scene: a correct calibration of the system allows to convert the 2D coordinates of recorded pixel in image plane to world 3D coordinates of the object points using triangulation rules, an vice versa. The aim of the calibration is therefore to find the correct geometric relationship describing this perspective projection and, therefore, the determination of the exterior orientation of both camera and projector and the interior parameters of the camera.

Considering such systems, it is convenient to model the projector as an inverse camera so that all the consideration about passive stereo systems are valid and the calibration rules can be directly and easily applied. In fact, until now, the camera calibration has been intensively studied and, as a result, a great number and variety of calibration algorithms exist. However, the projector model is normally a reduced model, where lens distortions are neglected or considered as second order variables.

Considering the most conventional calibration method, the approach requires the knowledge of the mathematical models governing the projector-camera pair and the identification of calibration artifacts. The standard sequence of steps involves the camera extrinsic matrix estimation (camera calibration), the estimation of the projection planes equation (projector calibration) and the assessment of the projective transformation (projector-camera calibration). It must be noticed that classic camera calibration relies on 2D image correspondence of 3D reference points, that depends on the projector calibration operation. In this way, the camera calibration affects the reliability of projector calibration degrading its accuracy (the projector calibration error can be 1 order of

magnitude higher than the camera calibration one). For this reason, other approaches have been studied, that propose the simultaneous calibration of the devices, in order to avoid the influences of the camera calibration on the projector calibration (Chen et al., 2009). The camera calibration algorithm proposed by Zhang (Zhang, 2000), for its simplicity and accuracy, is however the most used calibration strategy: the process uses the pinhole camera model and permits the simultaneous definition of intrinsic and extrinsic camera parameters, namely focal length, optics aberration, central point's coordinates, skew (IO), and position and orientation (EO). The approach consists in capturing images of a planar checkboard calibration artifact at several distances and orientation; the algorithm will then use relation between the checkboard corners coordinates and the world coordinates to calculate the camera calibration parameters.

Next paragraphs will deepen the calibration approach.

### 3.4.1 Calibration artifacts

The system calibration procedure requires a sufficient number of 3D references points, placed with a sufficient level of accuracy in the 3D space and having known WCS. This function is carried out by calibration artifacts of several shapes and dimension. In general, the function of the calibration artifacts is to provide enough reference data to be used into system calibration model for a correct 3D measurement and estimation of system parameters. According to their dimensions, the calibration artifacts can be classified four categories.

*0D approach (self-calibration)*: this technique does not require any calibration artifact. The input data for the camera calibration are the image point correspondences.

*1D approach*: a work proposed by the senior member of IEEE, Zhengyou Zhang (Zhang, 2004), illustrates for the first time the possibility to calibrate cameras using 1D objects, such as points aligned on a line. This technique has resulted fundamental for multiple cameras calibration, when calibration objects need to be visible simultaneously by cameras independently mounted.

*2D plane-based approach*: the most commonly employed calibration artifacts are 2D planes: this type of artifacts significantly differs from each other depending on the characterizing pattern and the material of the board. It can be a squared array or uniformly distributed circular marks printed on a rigid support (metallic, plexiglass, plastic etc.) (Figure 3.21). The most used 2D artifact is a planar bi-dimensional checkboard (Zhang, 2000; Anwar et al., 2010; Zappa et al., 2010; Moreno and Taubin, 2012;).

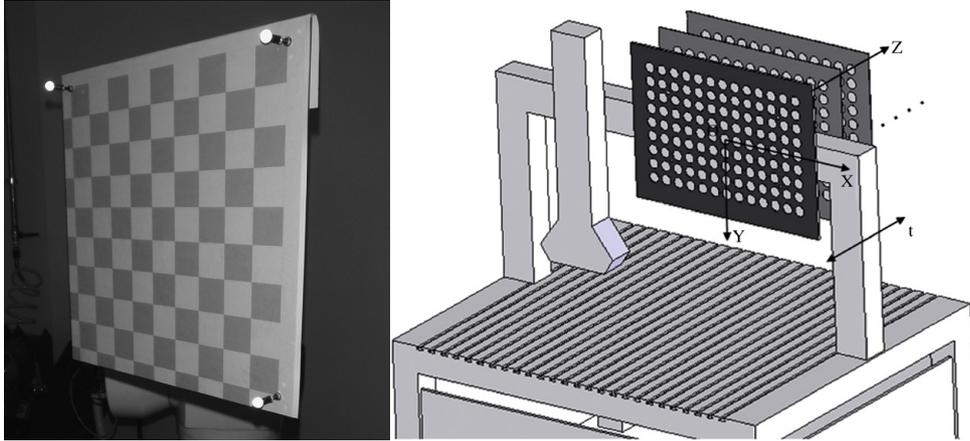


Figure 3.21. Calibration artifacts example: classic checkboard (left) (Zappa et al., 2011) and circular marks (right) (Chen et al., 2009).

*3D artifact approach:* less common but used for some experimentation in 3D imaging field is the employment of 3D artifacts: the object usually consists of two or three orthogonal planes (Figure 3.22) and the camera calibration is performed thanks to the knowledge of the 3D geometry with high precision.



Figure 3.22. 3D Calibration artifact composed by three orthogonal planes (Sturm and Maybank, 1999)

Beyond the artifact shown in Figure 3.22, can be considered a 3D approach also the technique that implies the translation of a plane in the space: considering 2D planes, during the calibration, the artifact is placed in various poses with respect to the system, and typically it is progressively translated and rotated away at fixed steps from the device.

### 3.4.2 Camera and projector models

Defining two coordinates systems, the world coordinates system  $(X, Y, Z)$  of origin  $O$  and the device coordinate system  $(x, y, z)$  of origin  $o$ , an object point  $P$  can be expressed in WCS and in camera image plane as follow:

$$P^W = [X \ Y \ Z]^T$$

$$P^c = [x \ y \ z]^T$$

The transformation from  $P^W$  (WCS) to  $P^c$  (camera coordinate) is a rigid body transformation. Considering a rotation matrix  $R^c$ , a translation matrix  $t^c$ , a scale factor  $s$  and the intrinsic parameter of the camera, the roto-translation between the WCS and the camera coordinate system can be expressed by

$$P^c = C^c P^W$$

$C^c$  is a 3 x 4 perspective projection matrix (PPM) of the form:

$$C^c = s M^c [R^c t^c]$$

The intrinsic parameters of the camera are the coordinates of the principal point  $c_x, c_y$ , and the components of the focal length along x and y,  $f_x, f_y$ . They are contained in  $M^c$

$$M^c = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}$$

Since the projector is comparable to an inverse camera that projects images instead of acquiring them, this perspective projection model can be applied also to the projector. The 2 x 3 PPM is

$$C^P = s M^P [R^P t^P]$$

where

$$M^P = \begin{pmatrix} f_p & 0 & x_p^0 \\ 0 & 0 & 1 \end{pmatrix}$$

Hence, the transformation from world coordinates to projector coordinates is given by:

$$P^P = C^P P^W$$

Here  $P^P = [x^P \ y^P \ z^P]^T$  is a 2D point in projector image plane and  $[R^P t^P]$  are the extrinsic parameters of the projector (rotation and translation matrix).

### 3.4.3 Calibration procedure

The calibration of a SLS is a process that consists in a series of successive procedures strictly related to each other (Figure 3.23). For the sake of brevity, here it will be illustrated a standard calibration procedure that uses the translation of a classical calibration checkboard, leaving out all the cases involving 1D and 2D artifacts. Moreover, we assume that the light projected is codified with Gray Code. The basic idea is to collect images of the illuminated translating planar checkboard and successively, to estimate the parameters of both camera and projector in the chosen camera model. The procedure starts with the preparation of the calibration board: the artifact must be placed within the working distance of the projector in a first orthogonal position. Follows the acquisition of a full illuminated scene to find the coordinates in the camera image plane of the checkboard corners.

The procedure continues with the projection of the codified pattern on the checkboard and the relative image capturing. This step is repeated for each pose of the calibration artifact until covering all the working volume. The positioning can be performed by mechanic automatic system or, in the simplest cases, by an operator that is called to manually and carefully shift the checkboard. The central problem of the projector calibration is the estimation of the checkboard corner location in projector pixel coordinates: in fact, although the projector is described by the same model of the camera, it does not capture any image on which search for corners coordinates. This lack of data, however, can be balanced with the information provided by the structured-light code, that sets a relation between projector and camera pixels. There exist several methods to find this correspondence. A recent study of the School of Engineering of the Brown University proposes to find and use local homographies for each checkboard corner in the camera image (Moreno and Taubin, 2012). The method overcomes the non-bijectionality of the decoding pattern function, that does not allow the direct translation of the checkboard points coordinates from camera to projector one.

In general, the projector calibration is performed in three steps:

- Decoding of structured-light pattern;
- Association of each camera pixel with a projector row and column;
- Estimation of checkboard corner location in projector pixel coordinates;

In conventional SLS calibration, the camera calibration follows a standard approach. However, over the past years, several calibration methods that use a pre-calibrated camera have been proposed (Kimura et al., 2007; Falcao et al., 2009) and even some that do not require any camera calibration or pattern (Anwar et al., 2012; Drarèni et al., 2009). Once found the intrinsic parameters of camera and projector, the exterior orientation is determined through the calibration of the system, namely a stereo calibration. Since the projector is modeled as a camera, the system can be led back to a camera-camera standard structure, and therefore the classic calibration process follows.

Known parameters	Unknowns
<ul style="list-style-type: none"> <li>- Camera and projector IO</li> <li>- World coordinates (camera coordinates)</li> <li>- Physical dimension of calibration checkboard</li> <li>- Checkboard corner coordinates in projector and image plane</li> </ul>	<ul style="list-style-type: none"> <li>- EO of projector in world coordinates</li> </ul>

Table 3.1. Calibration input (known parameters) and unknowns.

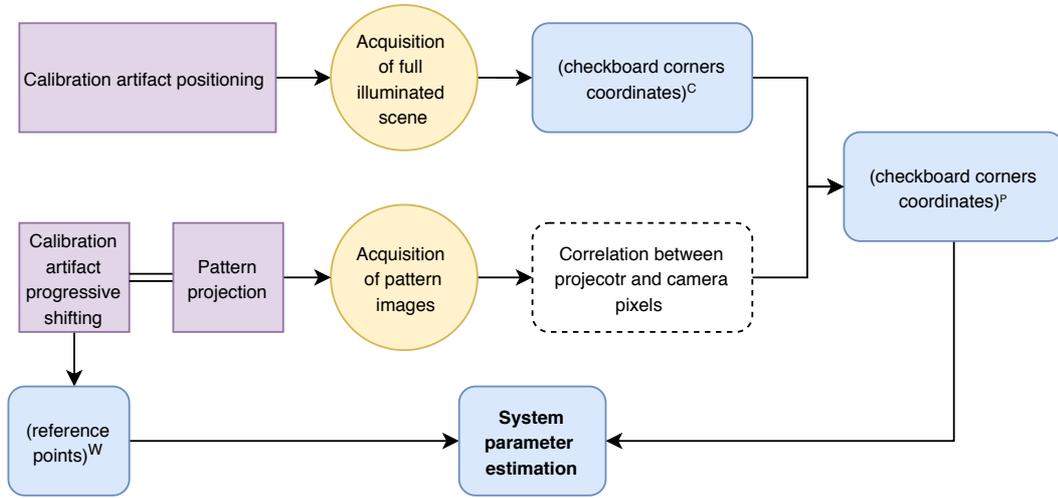


Figure 3.23. SLS calibration workflow

The accuracy of calibration is generally evaluated through the reprojection errors values. To ensure accuracy of the process, an objective function is used, that minimizes the sum of the reprojection errors related to the reference points onto projector and camera plane. The system parameters estimation is performed adopting the maximum likelihood approach based on the calibration model. If  $\Theta^C$  and  $\Theta^P$  are the set of respectively camera and projector parameters, and  $P^W(k)$  represents each reference point in WCS, the objective function can be express as follow:

$$\{\Theta^C, \Theta^P\} = \arg \min \sum_k \|g(P^W(k), \Theta^C) - P^C(k)\|^2 + \|g(P^W(k), \Theta^P) - P^C(k)\|^2$$

This optimization problem is non-linear and can be solved with several methods that provide an initial guess of projector and camera parameters. The solution is obtained by least-squares fitting method.

The system calibration itself, can be summarized as follow:

$$R^{PC} = R_P(R_C)^{-1}$$

$$T^{PC} = T^P - R^P (R^C)^{-1} T^C$$

where  $T^{PC}$  and  $R^{PC}$  represent respectively the translation and the rotation from the camera to the projector.

### 3.5 Artec Spider and Eva scanners

Recently developed, the Artec Spider and Eva are hand-guided scanners specifically designed for users who requires high precision 3D models. Relying on structured-light technology, they both are able to capture small- and medium-size objects shape with complex geometry without any contact with the surface.

Artec Spider, specifically, is an industrial high resolution 3D scanner characterized by metrological accuracy based on blue-light technology. The acquisition apparatus is composed by three cameras, a 6-LED flash and a texture camera, which acquires the chromatic information of the scene in real time during the survey (Figure 3.24). Its major advantages are reduced dimensions (190x140x130 mm) and lightweight (0,85 kg), together with high acquisition speed and resolution. These characteristics make the Spider scanner a valid support for industries that require the modeling of small, detailed objects with complex geometry, sharp edges and thin ribs. The fields of application of this device, however, are numerous and go from the reverse engineering to manufacturing, from biomedicine to cultural heritage. Table 3.2 contains its specification.



Figure 3.24. Artec Spider blue-light projection scanner (<https://www.artec3d.com>).

<b><i>3D resolution</i></b>	0,1 mm
<b><i>3D point accuracy</i></b>	0,05
<b><i>3D accuracy over distance</i></b>	0,3 mm over 1 m
<b><i>Working distance</i></b>	0,17-0,35 m
<b><i>3D reconstruction rate</i></b>	7,5 fps
<b><i>Data acquisition speed</i></b>	1 000 000 pts/s

Table 3.2. Artec Spider specifications.

Regarding Artec Eva, it is a structured-light scanner designed for the acquisition of medium-size objects (Figure 3.25). It is characterized by high-accuracy and resolution, as well as fast scanning and extreme portability. As Spider, Eva is lightweight (0,90 kg) and compact (262x158x63 mm). The large working distance, coupled with 0,2 mm resolution, allows to capture larger object maintaining excellent performances in terms of accuracy. Artec Eva is used mostly for reverse engineering, rapid prototyping and accurate survey of machine components and furniture, sculptures, human bodies, cars, motorcycles and in general of a wide selection of medium-size objects. See Table 3.3 for specifications.



Figure 3.25. Artec Eva scanner (<https://www.artec3d.com>).

<b><i>3D resolution</i></b>	0,2 mm
<b><i>3D point accuracy</i></b>	0,1
<b><i>3D accuracy over distance</i></b>	0,3 mm over 1 m
<b><i>Working distance</i></b>	0,4-1 m
<b><i>3D reconstruction rate</i></b>	7,5 fps
<b><i>Data acquisition speed</i></b>	18 000 000 pts/s

Table 3.3. Artec Eva specifications.

Both scanners are related to ArtecStudio software that perform all the modelling process, from data acquisition and processing, to 3D reconstruction, model editing and management. Except for the survey scale and resolution, both devices in general provide some fundamental advantages that ensure simplicity of scanning and high quality results:

- The possibility to perform the shape acquisition without the need of any target, thanks to the use of hybrid geometry and color tracking methods for data capturing and processing.

- The automatic acquisition of chromatic information during the scanning phase. The texture resolution is up to 1.3 megapixels, the algorithms include the color enhancement and automatic glare removal, resulting in vivid and realistic color 3D assets.
- Automated processing of 3D data through a batch processing.

### 3.5.1 Calibration

Artec Spider and Eva 3D scanners are delivered pre-calibrated: this means that the device is ready-to-use just after the activation, without the need of any other parameter adjustment. However, a correct calibration is recommended to avoid any parameters setting problems related to accidental drops and abrupt movements during transportation or handling. Therefore, is highly recommended to calibrate the scanner before every survey to restore the initial orientation parameters. *ArtecStudio* provides a special utility called *Diagnostic Tool*, which enables the user to perform the instrument calibration rapidly and easily. It is a universal tool devoted to the parameters reconstruction of both Spider and Eva; however, a slightly difference occurs between the two procedues.

The calibration of Artec Spider requires an equipment, which is provided coupled with the scanner itself, and consists in (Figure 3.26):

- a *calibration pattern* indicating the position to be assumed by the calibration artifact
- a *scanner stand*
- a *2D calibration artifact*, printed on plexiglass rigid support. The pattern consists in couples of points placed uniformly on the plane with varying orientation.

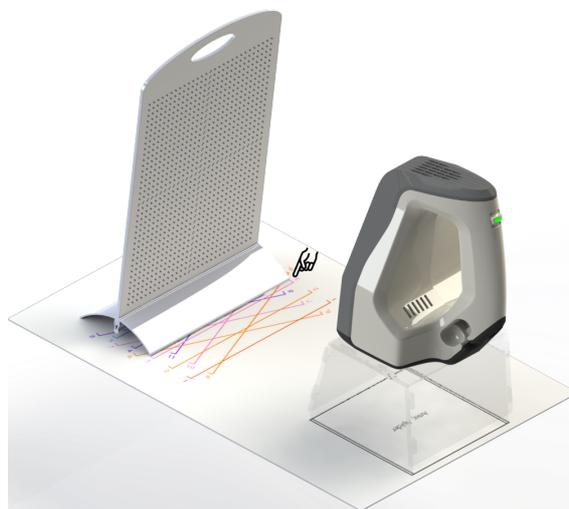


Figure 3.26. Calibration set: Spider scanner over the stand, positioned on the calibration pattern in front of the calibration artefact (ArtecStudio 15 Manual).

The calibration procedure is standard: the pattern assumes fifteen different positions (marked with coloured lines on the calibration pattern) until covering the entire working area, while the scanner is projecting the structured-light on it. For each position, the cameras acquire an image that is instantly used by the software to compute the orientation parameters. Specifically, the *Diagnostic Tool* displays two planes, a green plane captured using the original calibration data, and a red plane captured using corrected calibration data. The operator is called to slightly move the rig maintaining

the main position indicated by the colored lines until the two planes become coincident (Figure 3.27). At this step, the software calculates the corrective orientation parameters. The preceding step is repeated sequentially for all the fifteen positions.

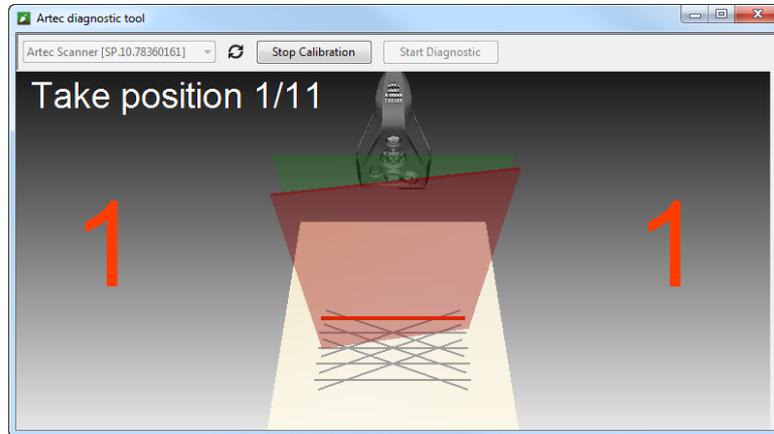


Figure 3.27. Diagnostic Tool during calibration: moving calibration rig to position number 1, green plane and red plane (ArtecStudio 15 Manual).

Considering Artec Eva, the calibration procedure does not involve the any calibration artefact and rig. The scanner is simply directed at the right angle to a flat, light, monochrome surface, such as a wall, from a distance of 650-700 mm. The surface is displayed as a blue plane, which represents the data captured using original calibration data. Once the scanner is activated, a yellow surface appears, corresponding to a surface captured using corrected calibration data (Figure 3.28).

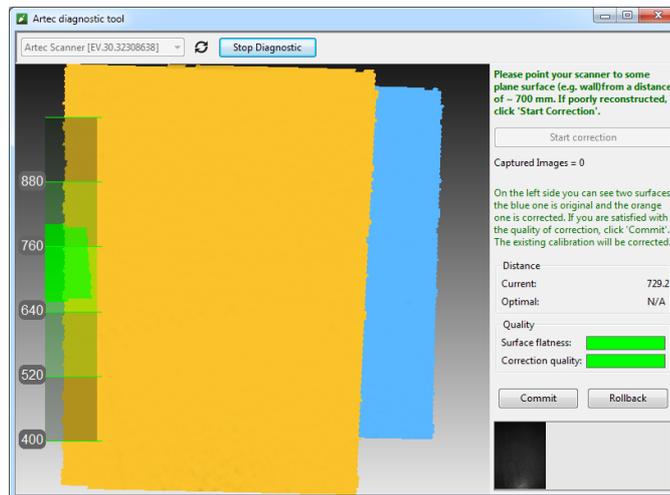


Figure 3.28. Artec Eva calibration process (ArtecStudio 15 Manual).

It is evident the deficiency of the calibration procedures described above. The parameter estimation conducted with the *Diagnostic Tool*, is completely up to the mathematical models and algorithms embedded in the ArtecStudio software which are not provided to the user. This results in the impossibility to set any calibration model and neither know which parameters are estimated and

how. These are common issues when dealing with active techniques, whose sensors are often sold as black boxes.

## 4. UNCONVENTIONAL APPLICATIONS OF HIGH-RESOLUTION ACQUISITION TECHNIQUES

### 4.1 Introduction

During the last decades, the attention for 3D range and imaging sensors has increased. The possibility to obtain high-resolution and metrically correct representation of different scale entities, have begun to interest a variety of application fields originally unrelated with these types of technologies, thus going beyond the fields for which they were initially projected, such as mechanical manufacturing industry, and surface quality control. Examples of fields that are approaching the high-resolution surveying methods are Geology, Civil Engineering, Medicine, Forensic, and Cultural Heritage, that take advantage of these techniques as integrated procedures.

Interesting applications in the field of Geology regard the capability of TLS to evaluate the surface morphology (or roughness) of natural materials (Mills and Fotopoulos, 2012), or the analysis of TLS point clouds for the determination of the failure geometry and extension of a landslide (Spreafico et al., 2016). In the field of Civil Engineering there are numerous possible applications: the widest spread ones are the structural monitoring (Bitelli et al., 2011; Mistretta et al., 2019; Artese and Zinno, 2020; Xu et al., 2018) and BIM/HBIM applications (Quattrini et al., 2015; Oreni et al., 2014); less conventional but still interesting is the employment of SLS for the surface model of road pavements (Bitelli et al., 2012).

The high-resolution surveying applications in the field of medicine started to be relevant already three decades ago, and today it is proved that 3D techniques, of both image- and range-based type, provide enormous benefits in medical care of patients. The advantages regard better surgical interventions planning, reduction of implantation and anesthesia time, and patient-specific solutions. A great variety of study cases have been presented by the scientific community, going from the 3D documentation of skin and bone injuries (Thali et al., 2003) to clinical bracket point registration (Nouri et al., 2015), and 3D printing in maxillofacial surgery, orthopedics, hearing aids (Kovacs et al., 2006; Heller et al., 2016). Also in forensic field, the help of high-resolution surveying methods has been realized in assorted fields: the use of laser scanner for forensic anthropology for example, help to quantify differences in skeletal features that are difficult to appreciate by eyes, thus improving the correct interpretation of human finds (Sholts et al., 2010; Perrone and Williams, 2019).

The reasons that have led such different fields to adopt 3D optical techniques more and more frequently are often shared. The foremost cause is the need of collecting data, of both metric and qualitative type, in order to create a database of digital information to be consulted and integrated over time. Together with the information collection, it is relevant to identify the accuracy of those data as concurrent cause for the great interest about the mentioned surveying techniques. Separate dissertations need to be done for image based and range based acquisition techniques.

Considering firstly photogrammetric reconstruction process, it is well known that the network design is a fundamental requisite to obtain overall quality, namely precision, reliability and economy, that however depend also on the parameters related to bundle adjustment, and the calibration strategy adopted (Fraser, 1997). Stated that the accuracy evaluation of photogrammetric reconstruction is not aside of a comparison with ground truth measurements (Nocerino et al., 2014; Guarnieri et al., 2004), it is reasonable to affirm that a close-range photogrammetry process, in presence of good network

geometry, can reach accuracy values proportional to Ground Sample Distance (GSD). This means the 3D reconstructions have a degree of resolution and good accuracy that make them capable to faithfully describe the world. The accuracy of range-based systems is well described in Chapter 3.

Commonly, according to datasheets provided by the manufacturers, TLS can reach an accuracy of few millimeters, in the range of 10-2mm, while a SLS accuracy is even higher, up to tenths of millimeters. As some works investigate, the real value is slightly lower, but still optimum (Chow et al., 2012; Kersten et al., 2008;). In general, the accuracy of the acquisition techniques influences the resolution of the survey, namely affects the quality of the reconstructed 3D models. High resolution acquisition therefore, means faithful, accurate, and metrically relevant reconstruction of the scene.

Besides being accurate and useful for documentation, the high-resolution acquisition techniques have the advantage of being also rapid and non-invasive; the instrumentation is versatile and user-friendly, especially during the last decade, when the technologies have made huge strides (Giancola et al., 2018).

### **4.2 High-resolution acquisition techniques for Cultural Heritage**

Cultural Heritage (CH) field has nowadays an increasing interest in 3D high-resolution acquisition methodologies for the digital recording of the cultural inheritance (Bonfanti et al., 2010; Francolini et al., 2019; Francolini et al., 2020). The spread of range- and image-based systems in this field is impressive and affects a wide range scale objects, from the large architectural structures, to very small archeological artefacts. As said before, the need for documentation is one of the most important reasons for the development of 3D surveying techniques in unconventional fields, and it is particularly true when speaking about World Cultural Heritage: a great part of objects belonging to it are constantly threatened by various natural or anthropic factors only partially counteractable (climate changes, natural disasters, wars, vandalism), and therefore are seriously in danger. For this reason, their recording is a fundamental operation, that aims to inheritance preservation and valorization (Remondino, 2011; Bitelli et al., 2017; Donadio et al., 2018; Aliberti and Picazo, 2019).

3D imaging and ranging techniques coming from Geomatic science field, due to their characteristics, allow to meet the need for Cultural Heritage knowledge, to be intended both as historic-artistic and both as geometrical-positional. The 3D products are accurate digital entities that constitute an important support to several aspects of the inheritance management: they are particularly effective in acquisition phase, being capable to record complex geometries in reduced time with high accuracy, but also in terms of data representation and visualization. In fact, the 3D digital products, generated in most of the cases through a photogrammetric reconstruction or with active triangulation sensors, allow an efficient communication of the inheritance, they improve their usability by the public and are fundamental for the promotion and divulgation of the CH (Dall'Asta, 2016; Bruno et al., 2010; Minto and Remondino, 2014). Moreover, the 3D high-resolution models considerably help the restoration process. It is well known that restoration is a complex task that involves multidisciplinary knowledge to accomplish various tasks, which go from the visual inspection and chemical study, to structural and image-based analysis.

High resolution 3D models give a concrete possibility to build an information system capable to support the storing, indexing, visualization and management of the information related to the restoration phases (Callieri et al., 2004; Girelli et al., 2020). A recently developed application is the physical scaled 3D reproduction of the objects: even though this technology is undergoing rapid

developments especially in industrial and medical field, 3D printed copies of archeological, artistic and architectures inheritance is nowadays a real possibility, which represent the fundament for new study approaches, cataloguing methods and visualization media (Neumuller et al., 2014; Balletti et al., 2017).

Over the last years interesting applications of high-resolution acquisition techniques for CH have been presented to the scientific community. Some of them aim to investigate the best practice for 3D survey in order to evaluate the quality of the reconstructed models depending on the acquisition procedure (Menna et al., 2016; Murtiyoso et al., 2019); other are mainly devoted to find new acquisition methods for efficient and effective acquisition (Kontogianni et al., 2017; Kontogianni et al., 2019; Francolini et al., 2020; Menna et al., 2017). Also, the issue of visualization improvement has been extensively treated (Dempski and Viale, 2005; Vergne et al., 2010; Defrasne, 2014; Carrero-Pazos and Espinosa-Espinosa, 2018). Several interesting publications about this theme focus on the enhancement of very small archeological finds characterized by complex geometry, namely objects bearing engravings or epigraphs (Hameeuw and Willems, 2011; Fisseler et al., 2014; Rothacker et al., 2015; Papadaki et al., 2015). All these peculiar applications take advantage of traditional surveying methods of Geomatic, namely image and range based sensors, for unconventional purposes.

Next paragraph will show some study cases concerning the issues just mentioned, focusing on the issues related to the acquisition phase with close-range photogrammetry and SLS, and the 3D model reconstruction process.

### **4.3 Longobard Basin of Santo Stefano Church, Bologna**

The field related to Cultural Heritage and the Geomatics area are particularly suitable for the development of applications with a strong interdisciplinary connotation. The work described here highlights the contribution of geomatic engineering to the medievalist historical research, where high-resolution acquisition is considered not only as a form of documentation and support for restoration but also as a tool for historical research and interpretation of the sources. The case study is the Longobard Basin, the so-called Pilate's basin, located in the Pilate's Courtyard inside the Santo Stefano Basilica in Bologna, Italy (Figure 4.1) (Borghi, 2010). The basin is a rare artifact which testimony the Longobard presence in Bologna. Its first function was to collect donations from the people to the Church of San Giovanni Battista, the site where it was originally located during the period of Liutprando (736-743) and Ildebrando (744): as a testimony of its function in that period, the basin bears an inscription on the upper edge. The epigraph has always been of interest to historians, evoking discussions regarding the interpretation of certain engraved words, which are not easily readable. This experimentation regards the application of 3D acquisition techniques on the basin as effective aid for its restoration process and historic analysis.

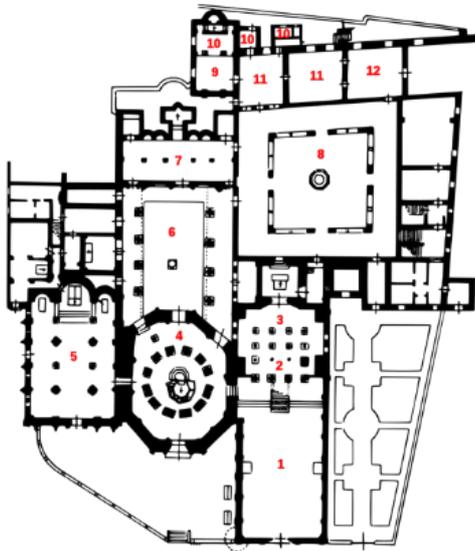


Figure 4.1. Santo Stefano Basilica (above), layout of the Basilica (below, left) and Pilate's basin (below, right); 1-3 Church of the Crucifix 2. Crypt 4. Church of the Holy Sepulchre 5. Church of Saints Vitale and Agricola 6. Pilate's courtyard 7. Church of the Trinity or the Martyrium, also called "Santa Croce" containing the ancient Nativity 8. Cloister 9-10-11-12. Chapel of the Bandage and museum.

### 4.3.1 The restoration

The conservation state of the artwork, despite various restoration interventions carried out over the years, was, until a short time ago, rather precarious. The reason for that resides in the fact that it has been exposed over the centuries to the atmospheric pollution, which has accelerated the physical mechanical decay of its glazing patina. Moreover, the exposure to the atmospheric agents as humidity and acid rains, has contributed to the formation of an environment which results suitable for biological proliferation, causing an additional alteration and deterioration of the stone material. For all these concurrent reasons, the superficial stone layer has been subjected to severe damages, which has led to the decision to restore it. The last restoration intervention has been carried out in 2017: it has acted on the numerous drip dark marks on the outside surface, on the patina covering its foot and the inside part of the basin.

The documentation of the restoring intervention has been carried out with three-dimensional acquisition of the basin, realized by close-range photogrammetry technique. The aim is to give a global 3D representation of the basin before and after the restoration phases, in order to create the first database of the basin's information. In this first phase, the realization of the 3D models has provided the possibility to conduct comparison analysis of the object before and after the intervention on the damaged surface.

The surface acquisition was performed using a full frame DSLR camera (Canon EOS 6D equipped with a wide-angle lens) and adopting a multi-photo approach: considering the survey before and after restoration, 63 and 175 photos have been taken respectively, assuming different positions around the basin (Figure 4.2). Since the basin is located outside, both surveys have been done without any artificial light source but taking advantage of the natural illumination. Coded targets were used for metrical dimensioning of the models. The photogrammetric software used for the surface reconstruction is Agisoft Metashape Pro. Figure 4.3 shows the reconstructed 3D models of the basin before and after the restoration:

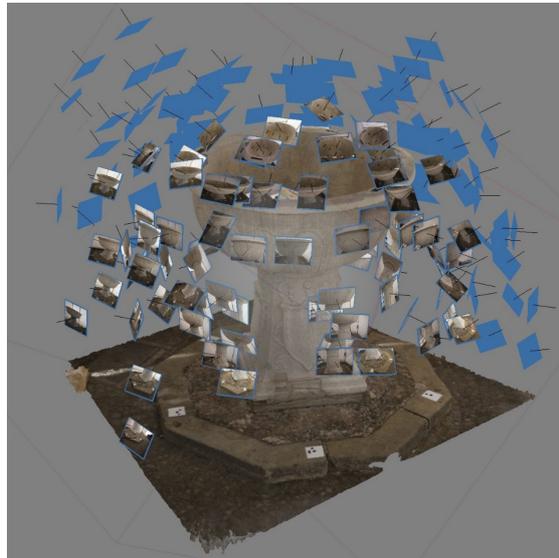


Figure 4.2. Photogrammetric acquisition of Pilate's basin. Multi-view configuration.



Figure 4.3. 3D textured models of the Pilate's basin before (left) and after (right) the restoration.

### 4.3.2 The historic interpretation

The history study of the Longobard basin concerns the historical placement of the object and the different interpretations that have been given to the inscription that is there reported. The basin inscription is composed by a series of signs engraved in the stone, thus, its study can be led back to the analysis of an epigraph. In general, in studying of epigraphs, it is required an accurate survey of the engraved surface in order to precisely detect the signs and to obtain accurate models capable to rigorously describe the surface.

Over the years, several surveying techniques have been developed, that mainly distinguish between contact and contactless. Among them, can be mentioned the direct drawing, making casts, enhanced photography (Weide et al. 1967), and Polynomial Texture Mapping (PTM), also known as Reflectance Transformation Imaging (RTI) (Earl et al. 2010); these classical methodologies, despite being widely used even today and being convenient for different reasons, do not provide satisfying results in terms of completeness, timing, usability and versatility (Francolini et al., 2018). This study takes advantage of the high-resolution acquisition technologies that permit to perform an accurate epigraph analysis exploiting the third dimension of the models.

The techniques involved are close-range digital photogrammetry and SLS. A preliminary testing phase is carried out to determine which of the two techniques could be considered the most appropriate for the high-resolution survey. The parameters evaluated for this purpose are:

- the capability of penetrating the engravings
- the resolution of the models

The test involves firstly the acquisition of a limited portion of the inscription using both techniques. The second step is the comparison between the two models through the calculation of the distance between them (referring to the same reference system) and the morphological analysis. Lastly, the resolution check.

#### **Photogrammetric acquisition and 3D reconstruction quality tests**

The close-range photogrammetric survey is performed in multi-photo approach, using a compact digital camera Panasonic, Lumix DMC-TZ60. The multi-photo approach involves 23 camera positions along the portion at a mean distance from the object surface of 11.4 cm; for constraints related to the dimension of the strip portion and the object itself, it was not possible to apply and take advantage of traditional coded targets; for this reason, 6 markers have been manually selected and positioned on the images through the photogrammetric software: their function consists in being reference points for the image matching procedure, aiming to facilitate the software in identifying homologous points on the images (Figure 4.4). The strategy adopted in positioning the markers consists in individuating three couples of points (pt.1\_pt.2, pt.3\_pt.4, pt.5\_pt.6) of known measured distance, at specified positions on the strip portion. In order to ensure a correct scaling of the final 3D model, the three distances have been introduced as scale bars, one for each couple of markers (Table 4.1).

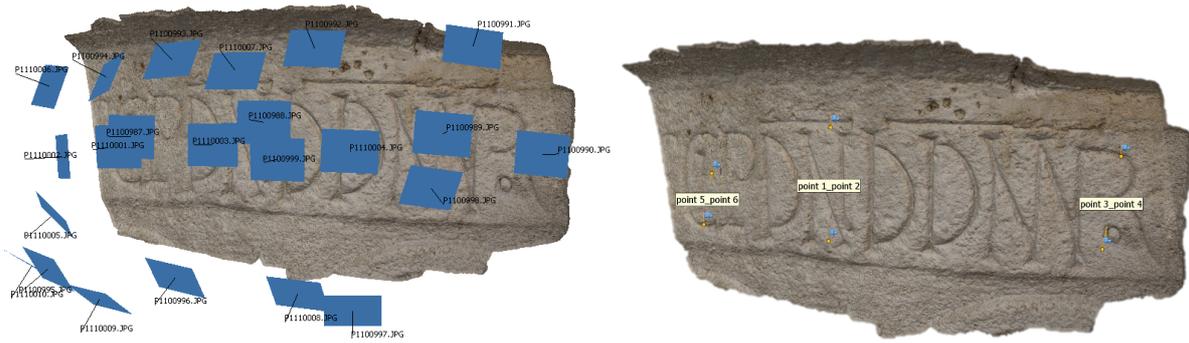


Figure 4.4 Photogrammetric survey of the portion on the inscription: 23 camera positions assumed during the survey (left), and the 6 markers (right).

<i>Scale bars</i>	<i>Distance (m)</i>	<i>Error (m)</i>
<i>pt.1_pt.2</i>	0.067	0.0011
<i>pt.3_pt.6</i>	0.049	-0.0003
<i>pt.5_pt.6</i>	0.033	-0.0015
<i>total</i>		<i>0.0011</i>

Table 4.1. Three scale bars adopted with the relative errors and total average error.

The performances of the photogrammetric reconstruction have been tested varying the presence of markers and varying the accuracy of the camera position estimates. Considering firstly the employment of markers, interesting considerations can be made. The reconstruction test is performed using the high accuracy mode during both photo orientation and dense cloud reconstruction, which is the recommended solutions to obtain accurate camera position estimates.

Regarding the photo orientation, the Table 4.2 shows that in both cases all the 23 cameras are aligned within a matching time of 2 minutes and 17 seconds without the use of markers, and 3 minutes and 36 seconds with the markers; and an orientation time of respectively 47 seconds and 8 seconds. In the first case the recognized tie points are less numerous (21.659) than in the case that considers the markers (32.051), and the max reprojection error over every tie point shows the same behavior; it is lower for the reconstruction with markers, reaching a value of 2.60 pix (with respect to 13.70 pix for the reconstruction without markers). The results confirm the good influence of markers on the photogrammetric reconstruction in terms of orientation quality and time, at the expense of matching time that, however is contained in the order of few minutes.

The final models' characteristics are summarized in Table 4.3.

	<i>No markers</i>	<i>6 markers</i>
<b><i>Cameras aligned</i></b>	23/23	23/23
<b><i>RMS reprojection error</i></b>	1.37 pix	0.50 pix
<b><i>Max Reprojection error</i></b>	13.70 pix	2.60 pix
<b><i>Tie points</i></b>	21.659	32.051
<b><i>Matching time</i></b>	2 min 17 sec	3 min 36 sec
<b><i>Orientation time</i></b>	47 sec	8 sec

Table 4.2. Comparison between photo orientation taking advantage of 6 markers and without them.

	<i>No makers</i>	<i>6 markers</i>
<b><i>Dense cloud point count</i></b>	17.962.469	18.410.987
<b><i>Mesh face count</i></b>	3.585.257	3.682.172
<b><i>Overall processing time</i></b>	41 min 4 sec	38 min 28 sec

Table 4.3. Final models' characteristics: dense point cloud, mesh and processing time.

This preliminary test has led to choose the 6-markers orientation solution for the successive test on the accuracy of the camera position estimation setting; despite being almost equivalent in terms of processing time and mesh degree of detail, the results in terms of reprojection errors are more satisfying. Within this analysis, highest accuracy mode effects are studied and compared with the high accuracy results obtained yet, taking into consideration both photo orientation and dense cloud reconstruction; since the highest resolution mode is called to upscales the images by factor of 4, it is expected to improve the results in terms of resolution, issue that is of primary importance for small objects reconstruction.

Table 4.4 summarize the orientation and surface reconstruction results employing the highest accuracy orientation. A first immediate consideration must be done about the processing time (Figure 4.5) starting already from the orientation time: in highest accuracy mode, this process lasts almost 51 minutes, which is a considerable amount of time if compared to the high-accuracy modality. The same consideration is valid for all the other processing, which result to be much more time-spending, making the overall processing time of highest accuracy mode more than 3 times longer than the high accuracy one. The processing time is a factor that influences the overall process; specifically, the high value of processing time in highest accuracy mode can impact considerably on the employment of the photogrammetric reconstruction, especially if the requirement of 3D models fall outside the academic research where the requirement for good results in reasonable time is an important factor.

<i>Highest accuracy</i>	
<i>Cameras aligned</i>	23/23
<i>RMS reprojection error</i>	0.64 pix
<i>Max Reprojection error</i>	4.68 pix
<i>Tie points</i>	37.894
<i>Matching time</i>	50 min 59 sec
<i>Orientation time</i>	1 min
<i>Dense point cloud</i>	67.064.058
<i>Mesh face count</i>	13.391.927
<i>Overall processing time</i>	2 h 31 min 4 sec

Table 4.4. Alignment and surface reconstruction results employing the highest accuracy orientation.

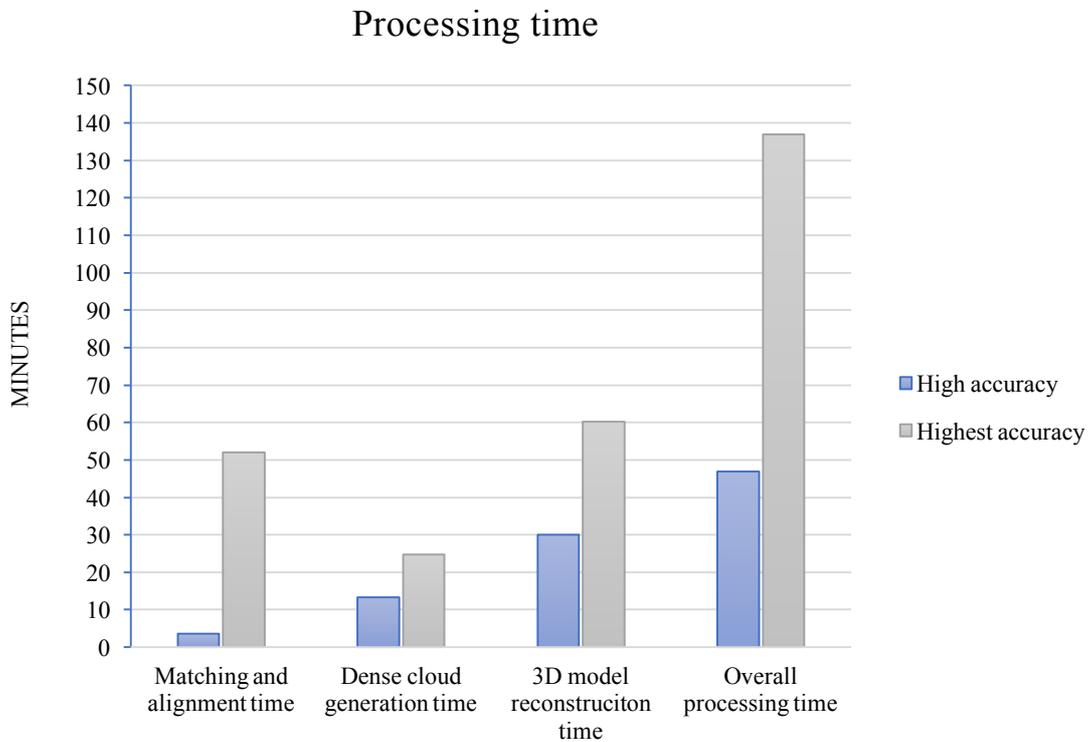


Figure 4.5. Comparison between the processing time in high resolution and highest resolution mode.

However, considering the dense point cloud and the final mesh reconstruction, the advantage of highest accuracy mode come up overwhelmingly. The point cloud generated in highest accuracy mode is composed by 67.064.059 points, over 18.410.987 points of high accuracy orientation, coming up in meshes characterized by 13.391.927 and 3.682.172 faces respectively. For the reasons listed above, the 3D reconstruction of the inscription is performed in highest accuracy mode. The final 3D model is shown in Figure 4.6



Figure 4.6. Final selected model reconstructed with highest accuracy settings.

#### **SLS acquisition and 3D model reconstruction**

The SLS acquisition of the portion of the inscription is performed with Artec Spider from Artec3D. The survey focuses on the engravings and thus the scanning has been performed carefully, changing the inclination of the device in such a way that the small details and the edges of the inscription could be acquired completely without any lack of data or holes. One scan has been sufficient to well describe the small portion surface, for a total number of frames of 682 and a very reduced acquisition time, in the order of few seconds. The first model after the first rough automatic alignment of the frames is shown in Figure 4.7.

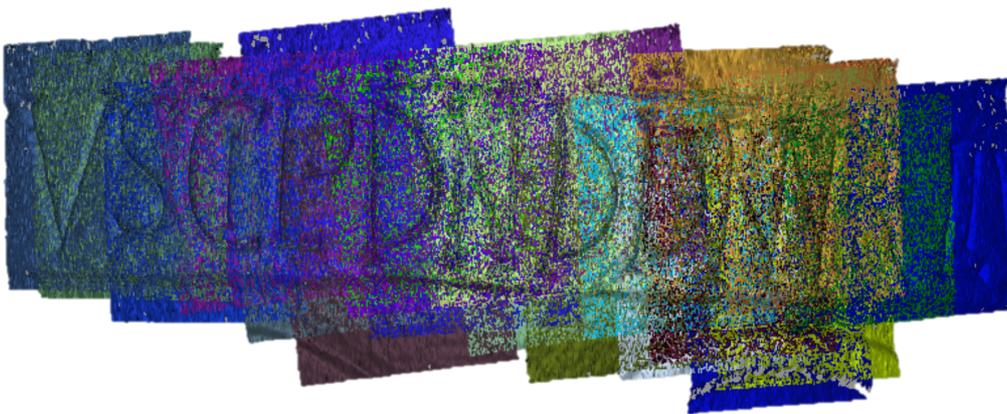


Figure 4.7. Rough model of the inscription portion. Different color represents the frames composing the scan.

The post-processing of the data involves the global alignment improvement that refines the search for correspondences between frames, the removal of outliers and the polygonal 3D model

reconstruction: this last operation is performed in sharp mode, whose algorithms can accurately reconstruct fine features and edges with a high level of detail. For Spider scanner, the highest achievable model resolution in sharp fusion mode is 0.10 mm; the final three-dimensional model of the inscription, has been built with a set resolution of 0.3 mm, which is comparable to those achieved with the close-range photogrammetry.

In order to give a radiometric content to each polygon composing the mesh, texturing step applies. The texturing, or texture mapping, is based on the color information of frames acquired in real-time during the shape acquisition: the method cuts the mesh surface in polygons and unfolds them so that they can fit the size of the images containing the radiometric information.

The final model is shown in Figure 4.8. However, it does not contain any radiometric information for a common problem related to the texturing operation: the issue is related to the impossibility to apply a high resolution texture to a higher resolution model, problem that can be faced by simplifying the models, i.e. reducing their resolution. In this case, texture information is not necessary for the experimentation purposes, so the model is visualized with arbitrary colors.



Figure 4.8. Final mesh of the portion of the inscription (false color).

### **Models comparison**

The comparison between photogrammetric and SLS models is not aside of their alignment and referencing in a common local reference system. Firstly, a rough alignment is manually imposed using the local based alignment on picked points of the 3D models. Stating that bending effects occur at the extremities of the photogrammetric models (Nocerino et al., 2014) and, for the geometry of the object, they couldn't be easily avoided, only the central part of the portion has been taken into consideration during the manual-picking point alignment. Then, the ICP algorithm is used to minimize the spatial difference between two data sets, basing on geometric feature information with the lowest possible amount of deviation in overlapping areas (Figure 4.9).



Figure 4.9. SLS and Photogrammetric models' alignment by picking-points. The points are individuated and selected manually, taking advantage of the surface peculiarities (roughness of the surfaces, engravings etc.)

Besides the well-known bending effects at the extremities of planar and non-closed surface reconstructed with photogrammetry process, also SLS is suspected to show analogous problem related to the extremities of long flat surfaces models; for this reason, an evaluation analysis of these errors is still complicated. From a first analysis of the aligned models, the measured relative maximum distance between them in correspondence of the extremities corresponds to 21 mm: this value, as anticipated before, is influenced by both bending effects related to SLS and photogrammetric reconstruction. Once referenced the models in the same system, the analysis begins with the extraction of longitudinal section profiles of aligned models. The profiles are obtained through the intersection of a plane  $xy$  perpendicular to the inscription surface plane (Figure 4.10).

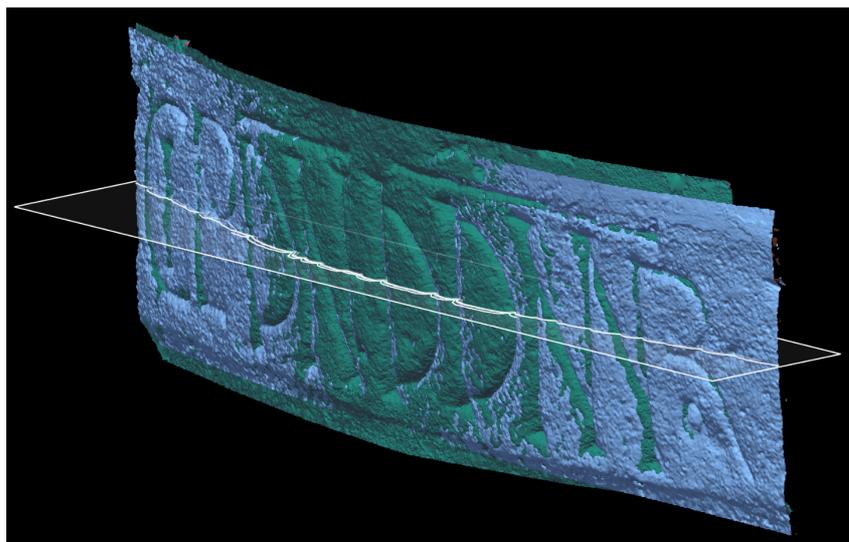


Figure 4.10. Plane and section profile intersecting the SLS and photogrammetric aligned models.

The measurement of the distance between the models in correspondence to the incisions follows. The highest measured distance achieved in the whole area considered in correspondence of the engravings is of 0.69 mm. (Figure 4.11). Considering that orange line represents the SLS profile,

can be stated that the 3D scanner is more capable to describe in deep the engravings than close-range photogrammetry.

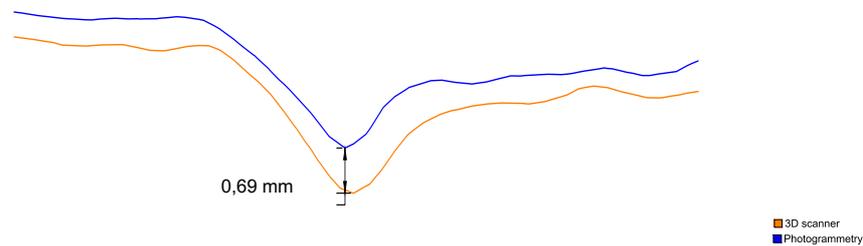


Figure. 4.11 Polylines representing the longitudinal section profiles of the portion of the inscription considered for the analysis (distance in mm).

A further analysis of the incisions can be made considering the 3D models of the portion of inscription as DEMs. The analysis on the incisions depth then, can be led back to a morphological terrain analysis where the incisions representing the inscription are comparable to riverbeds and the engraved surface to the terrain. The study has been performed using open software and involves morphometric indices commonly used in hydrology field, such as:

- Valley depth
- Local minima calculation
- Convergence index
- Topographic position index

### Valley depth analysis

The analysis involves the calculation of the difference between the elevation values of DEM and the ridge level. The ridge level is obtained using an algorithm that performs successively the individuation of the ridge cells using Strahler stream order on the inverted DEM and the interpolation of the ridge level (Strahler, 1957). The final step involves the subtraction of the original elevations from the interpolated ridge level. This analysis is performed on DEMs of the portion of the inscription obtained with photogrammetry and SLS technology. The results are shown in Figure 4.12: the analysis shows that in presence of narrow incisions SLS technology is capable to well describe the depth of the sign, while the photogrammetric reconstruction has lacks of sharpness and tends to flatten the three-dimensionality of the model. This consideration is valid for depth from 4 mm to 3 mm. Above these values, the two technologies have a similar behavior and reconstitute metrically comparable entities. The mean valley depth corresponds to 1,21 mm for the SLS model, and to 1,10 for the photogrammetric reconstruction, with a difference in the order of tenths of millimeters.

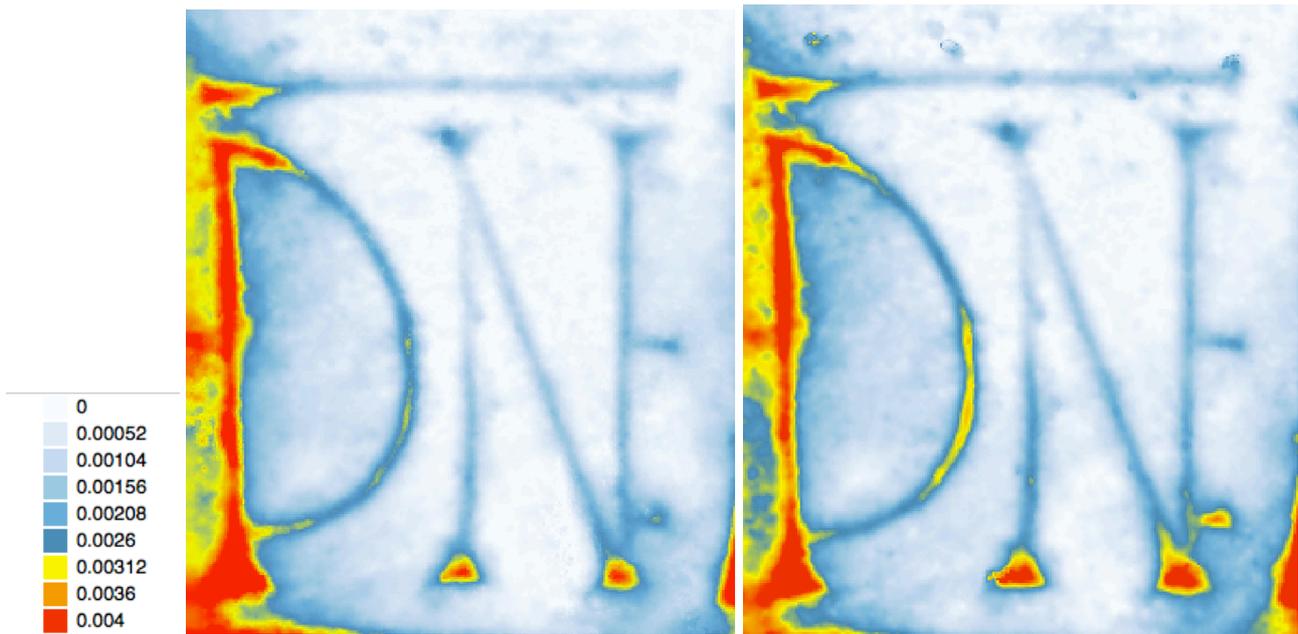


Figure 4.12. Valley Depth photogrammetry (left) and Artec (right). Values reported in legend in meters.

#### Local minima calculation

An interesting result about the capability of reaching the engravings in depth has been obtained through the calculation of local minima in correspondence to the inscription engravings. The Figure 4.13 shows the graphical distribution of the local minima along the Y direction: it can be noticed that, even though the trend is similar for both models, the values of local minima related to the photogrammetric reconstruction are lower than the local minima spotted on the SLS model (Figure 4.14).

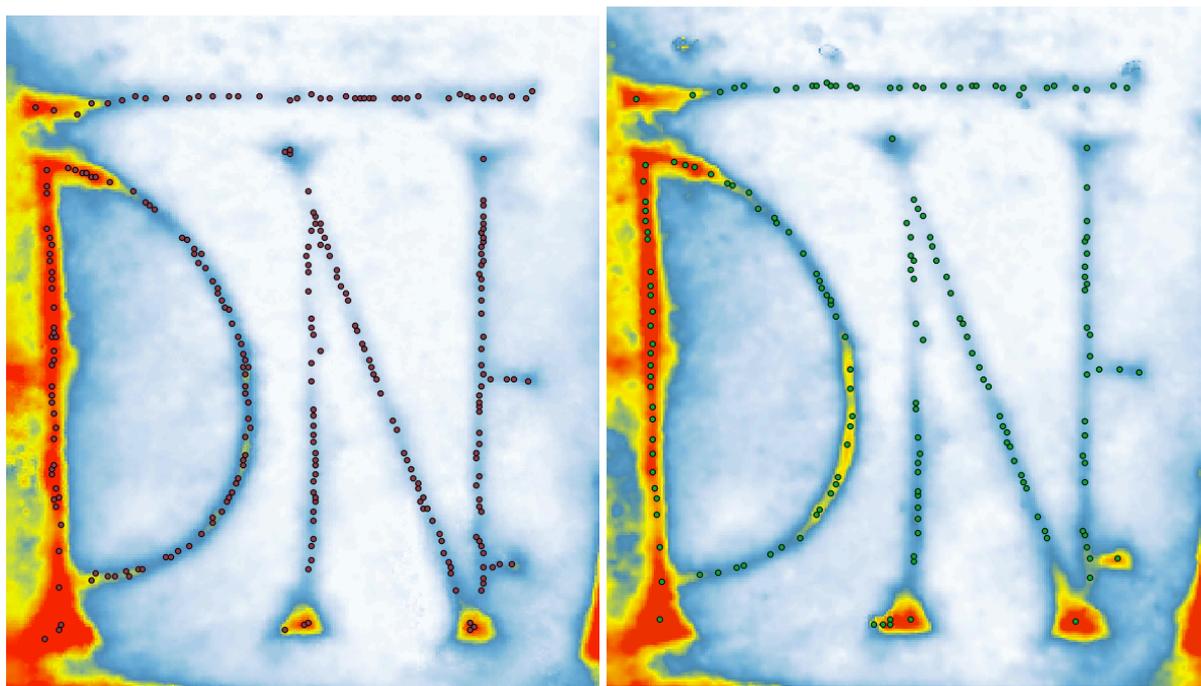


Figure 4.13. Local Minima representation: red and green dots on the valley depth maps corresponding respectively to photogrammetry (left) and SLS (right).

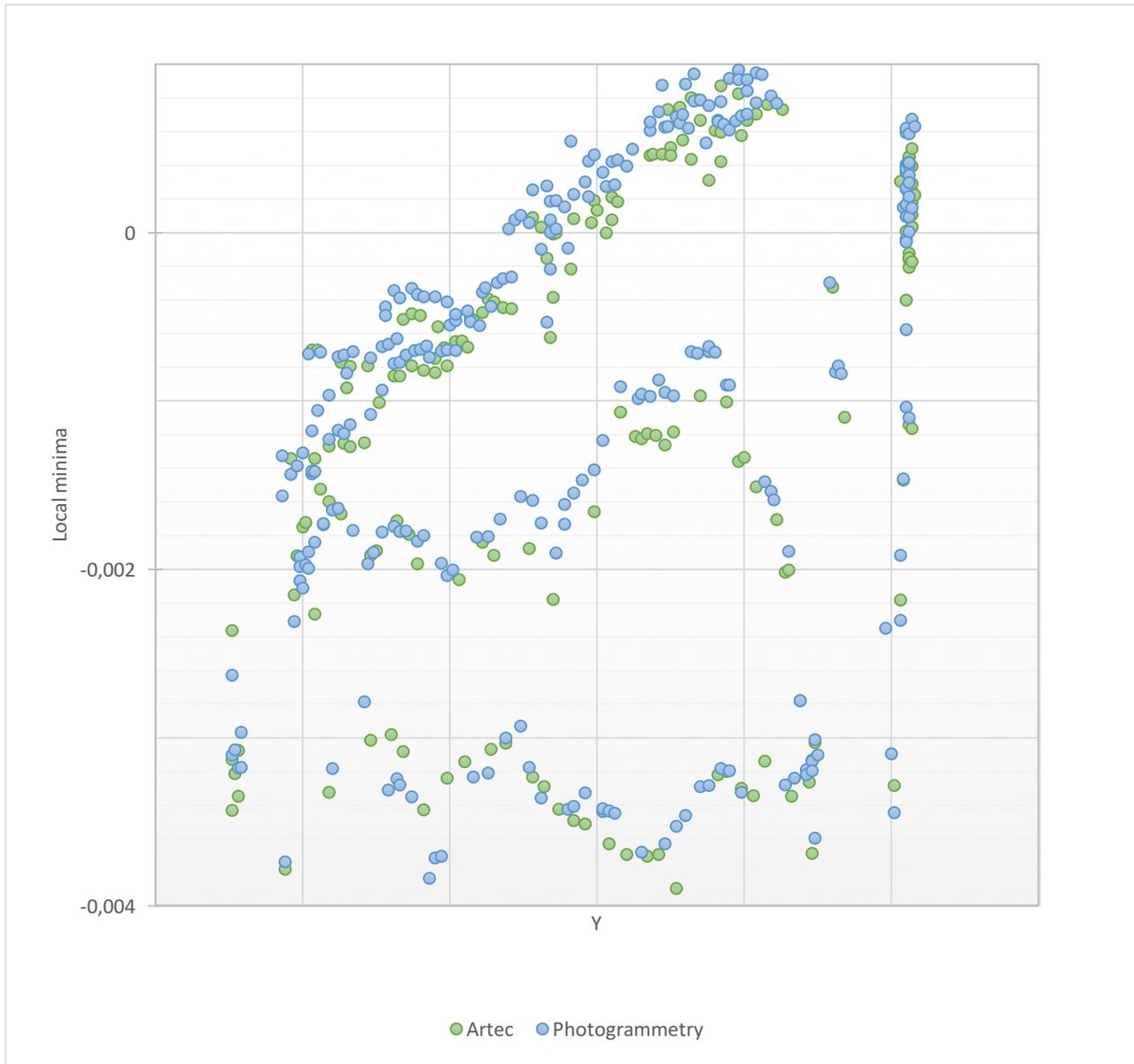


Figure 4.14. Local minima distribution along axis Y.

### Convergence Index

The calculation of convergence index allows to represent the morphological structure of a relief as a set of channels and ridges, namely convergent and divergent areas. This parameter symbolizes specifically the degree of agreement between the aspect direction of surrounding cells with the theoretical matrix direction, and is calculated as the mean aspect difference between real aspect and theoretical maximum divergent direction matrix representing ideal peak (Kiss, 2004). In the case of engraved inscription, the convergence index maps are represented as follows (Figure 4.15). The two maps are similar in aspect: the signs representing two letters result to be well described by the highest values of convergence (in blue), signifying that the algorithm have correctly individuated the engravings as convergent entities comparable to channels. Even though form maps it is difficult to appreciate the difference in convergence between the two models, the mean value of convergence for the SLS model is 0,171, while correspond to 0,138 for the photogrammetric model, meaning that the scanner acquisition has provided the reconstruction of sharper engravings.

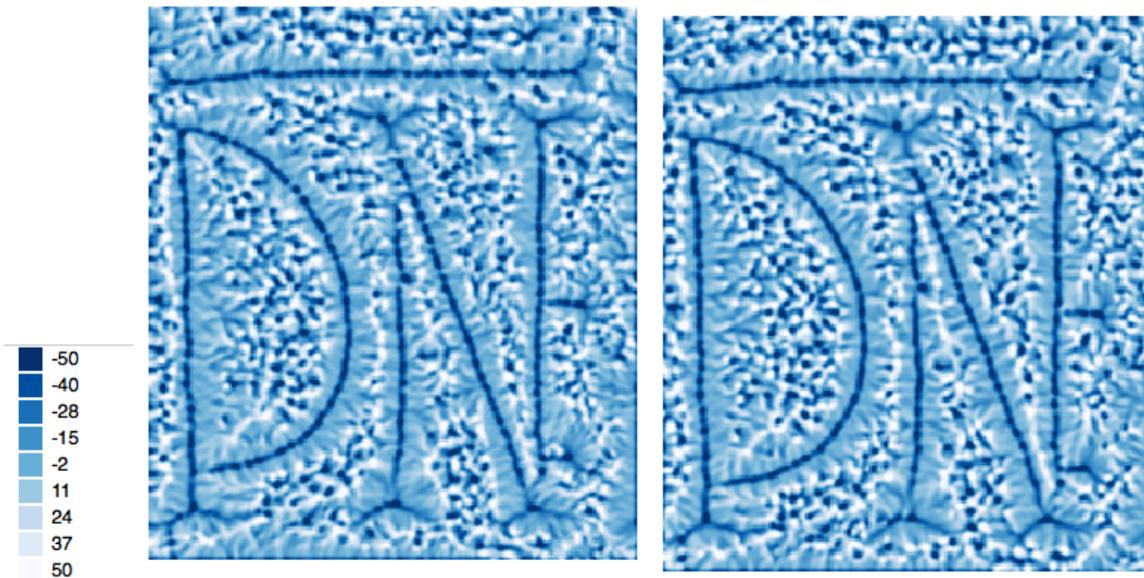


Figure 4.15. convergence index maps: photogrammetry(left) and Artec (right).

**Topographic Position Index (TPI)**

This morphological index is capable to represent locations that are higher and lower than the average of their surroundings, namely the ridges (positive values of TPI) and valleys (negative values of TPI). Its calculation involves the comparison of the elevation of each cell in a DEM to the mean elevation of a specified neighborhood around that cell (Figure 4.16). As a matter of fact, the topographic position is a scale dependent phenomenon (Guisan et al. 1999; Mokarram et al., 2015).

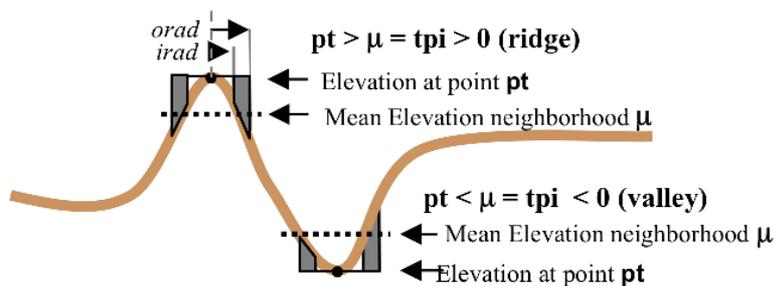


Figure 4.16. TPI calculation principle (<http://www.jenessent.com>).

In the case of the inscription, the TPI has been calculated separately for the photogrammetric and SLS DEMs, using three scales going from 1 to 8. The results show that Artec DEM is characterized by lower values of TPI in correspondence to the engravings than the photogrammetric DEM (Figure 4.17 and Table).

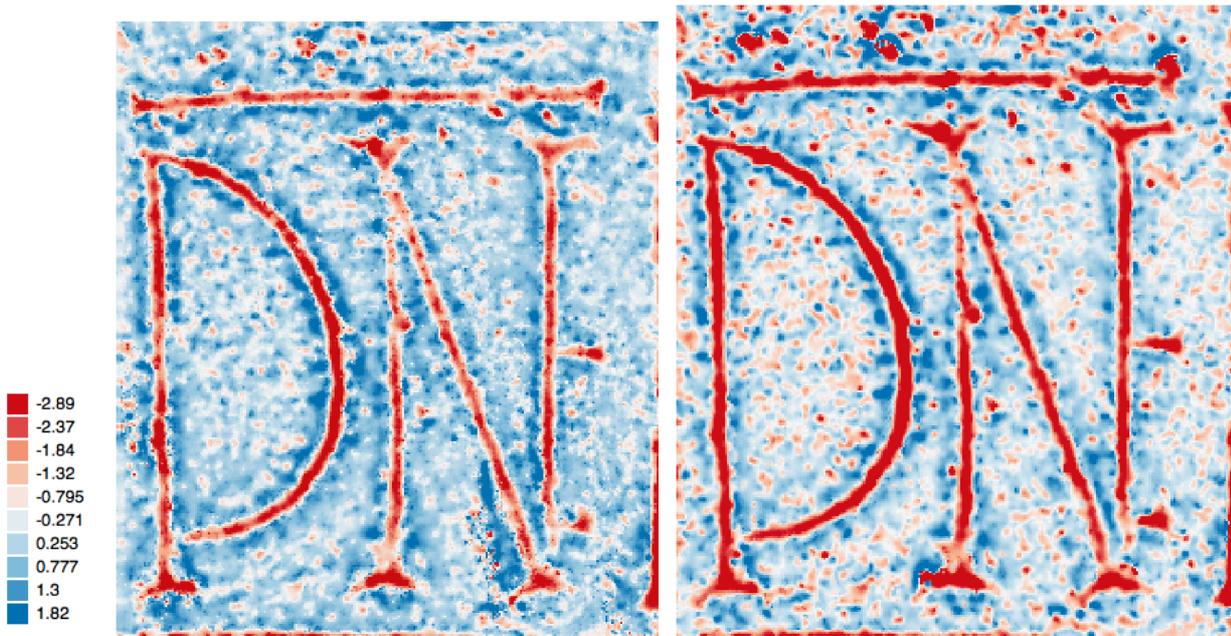


Figure 4.17. TPI maps: on the left, the photogrammetry map, on the right the SLS one.

### *TPI*

	<b>Photogrammetric DEM</b>	<b>Artec DEM</b>
<b><i>Minimum value</i></b>	-7.3951482772827	-15.15620136261
<b><i>Mean value</i></b>	0.0092871172136785	0.0089969926239213

Table 4.5. Minimum value of TPI for photogrammetric and Artec DEM

The results obtained from the morphological analysis satisfy the initial requirement of inscription depth analysis, providing synthetic indices that describe analytically and coherently the characteristics of the engravings. This method therefore represents a valid strategy for the interpretation of three-dimensional data of this type. The indices that have been proven to be most appropriate for the depth analysis of the engravings, are the *valley depth*, the local minima estimation and the *TPI*. On the other hand, the *convergence index* can be the starting point for the extraction of the signs, since it emphasizes well the convergent areas, namely the sharper marks of the inscription.

### **Models' resolution analysis**

Few considerations about the model resolution are necessary. Considering the model built through the photogrammetric reconstruction via Agisoft Photoscan, the value representing the model resolution is the Ground Sample Distance (GSD), or Ground Resolution, that here results in 0,3 mm/pix. This is a highly satisfactory resolution value in the field of close-range photogrammetry, and allow to consider this acquisition technique a valid solution for these types of application. The 3D model obtained from the structured light scanning and built in post-processing with Artec3D software, has been built with a resolution of 0,3 mm, a value that has allowed the models to be

rigorously and easily compared. This value is attributable to the mean inter-distance of the vertices composing the model, that is a parameter to be selected by the operator during the last step of the data processing, namely the mesh reconstruction. This value, even though being sufficiently high for high-resolution applications, however can be further improved to reach values up to 0,1 mm. This extremely high resolution value ensures very detailed 3D models, and is especially convenient in case of rough surface reach of small details; on the other side, it does not allow the application of satisfying resolution texture and its elaboration requires elevated computational power.

Considering the results of the morphological analysis just described, the considerations about the section profiles extraction, and the resolution analysis, can be stated that SLS technique is more suitable for the acquisition of the inscription because is more capable to acquire and represent the engravings. The SLS model of the whole circular strip containing the inscription of interest, is represented in Figure 4.18: it is composed by 1500000 vertices and has a resolution of 0,1 mm (Table 4.6), value that ensures an extremely high degree of detail and let the object to constitute a fundamental support to the analysis of the engraved text.



Figure 4.18. Final 3D model of the ring containing the inscription

<b><i>Model face count</i></b>	21.975.927
<b><i>Vertices</i></b>	11.053.903
<b><i>Resolution</i></b>	0.1 mm

Table 4.6. Final model characteristics

The model traces accurately the circular envelope of the engraved strip, causing however a visualization problem related to the legibility of the text. The circular shape of the model, indeed, does not allow an easy and rapid visualization of the text, unless rotating the model during the reading. The approach chosen to solve this issue is the digital unrolling of the surface. This operation involves digital procedures that approximates the model to a regular cylinder of height 0.11 m and radius of 0.61 m, and then projects it on a fictitious plane in the three-dimensional environment. The result is an entity, still three-dimensional, corresponding to the flat envelope of the inscription (Figure 4.19).

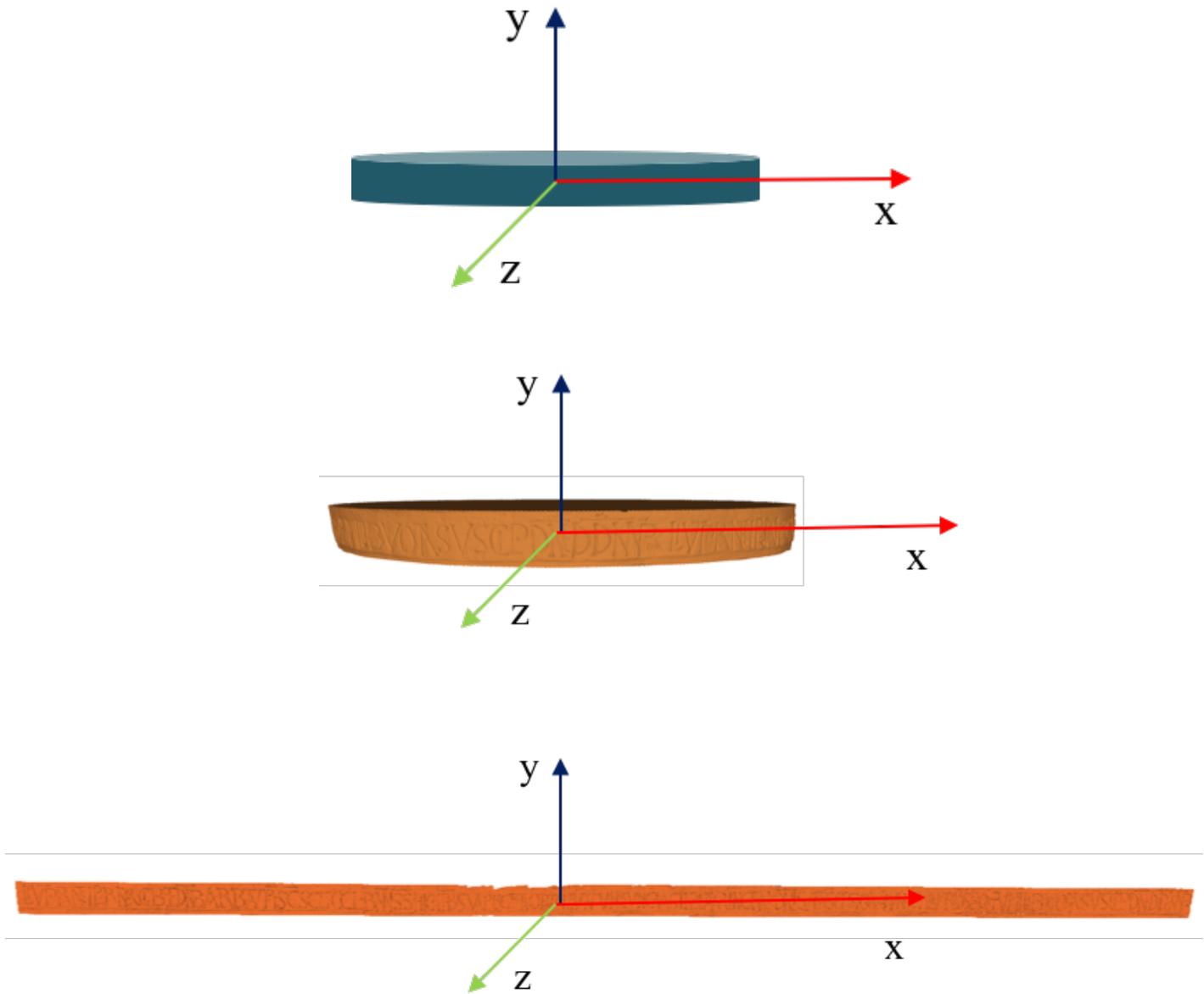


Figure 4.19. Unrolling operation: the digital operation approximates the model to a regular geometric solid, in this case a cylinder, of known dimensions and orientation in the space (left). The geometric solid surface is projected on flat fictitious support (xy plane) in the three-dimensional environment along a specified axis, in this case x axis.

Even though this digital solution provides a satisfactory improvement of the text readability, the three-dimensionality of the signs can be further enhanced to meliorate the overall result and allow an easier analysis of the text. For a better visualization of the model, a shading algorithm is used: the Lambertian Radiance Scaling. It is developed to allow the enhancement of concavities (or convexities) of a surface under arbitrary illumination, thus being capable to emphasize the engravings of a surface: the principle relies on the reflected light intensity adjustment, which depends on surface

curvature and material characteristics. The quantity considered is therefore the radiant energy of the object, namely the outgoing radiance on an object surface. The reflected radiance equation assumes the following shape:

$$L'(p \rightarrow e) = \int_{\Omega} \rho(e, \ell) (n \cdot \ell) \sigma(p, e, \ell) L(p \leftarrow \ell) d\ell$$

The mathematical quantities involved are: the enhanced radiance  $L'$ , a surface point  $p$ , the direction toward the eye  $e$ , the surface normal  $n$  at point  $p$ , is the hemisphere of directions around the normal  $\Omega$ , the light direction  $\ell$ , the material bidirectional reflection distribution function (BRDF)  $\rho$ , a scaling function  $\sigma$  and the incoming radiance  $L$ . (Cignoni et al., 2008; Vergne et al., 2010; Bitelli et al., 2014; Francolini et al., 2018). Moreover, artificial lighting in Meshlab ambient is applied, to enlighten part of the text of particular interest applying and varying the direction of the light beams (Figure 4.20)

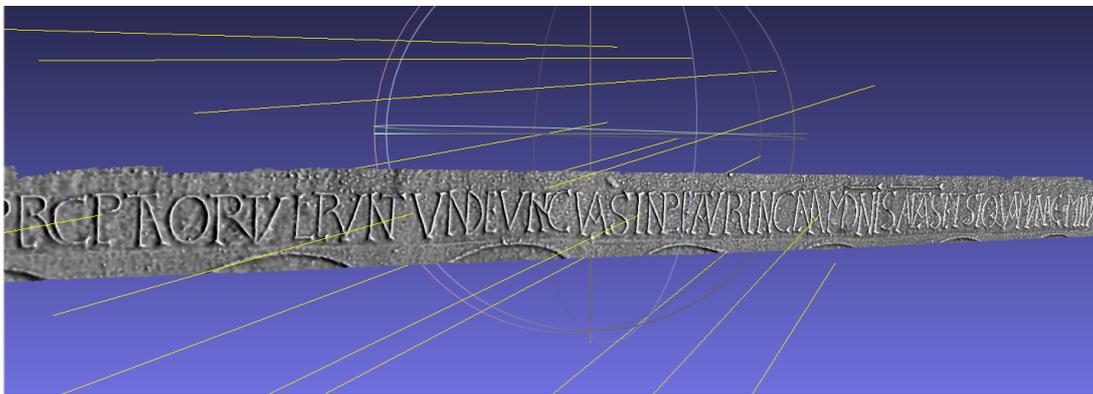


Figure 4.20. Artificial illumination in Meshlab Ambient. In yellow, the varying direction of the light beams.

The part of the text that interests the historians is the acronym reported in Figure 4.21. The interpretation of these signs is subject of numerous debates that regard specifically the last letter. Before 1800, what in the previous version of acronym has been written as B, has been successively taken as L, LR, and ER. The respective interpretations follow: *Johanni Baptiste* (IHB), in *honorem loci* (IHL), in *honorem loci religiosi* (IHLR), *inhabitants* (IHB), and *Iherusalem* (IHER). During the nineteenth century, another interpretation has been hypothesized by the scholar Gallinetti, that read the acronym as *in homorem (ecclesiae) pro luminaribus* (IHPL) (Gallinetti, 1985). Finally, in 2007, some new points of discussion have raised, claiming that the acronym can be interpreted as *In Hac Basilica* (IHB). This last hypothesis has been accepted by the philological community from an interpretative point of view, which is philologically coherent (Cosentino, 2007). However, in 2010, a new research on the possible interpretations of the sign has been carried out (Borghi, 2010), hypothesizing a further meaning of the acronym, *Iherusalem* (IHR). This hypothesis would induce to consider its connection with the symbolism of Jerusalem related to the complex of Santo Stefano, which is thought to date back later than the original Longobard artifice. Furthermore, this theory would suggest that, after the Longobard presence, the outline of the letter R was probably subjected to modifications, leading to the transformation in a B.

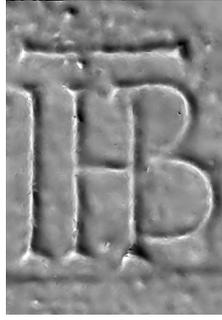


Figure. 4.21. Studied acronym, composed by three letters. The last one is subject to numerous debated on its interpretation.

For the mentioned reasons, the 3D model analysis mainly involves the last letter. In order to evaluate the affinity of the debated letter with a B or an R, a depth analysis of the engraved signs has been performed. Two generic and representative “B” and “R” are taken from the inscription and considered as reference letters; then, the last letter belonging to the acronym is compared to these two in order to individuate differences or similarities. The depth analysis consists in digitally reconstructing a plane tangential to each of the three considered letters, and successively generating parallel planes with an offset of 0,5 mm in depth (Figure 4.22). This operation permits to highlight the details of the incision in deep, which at a superficial sight are difficult to appreciate. Lastly, the section profiles are extracted as polylines in CAD environment and compared.

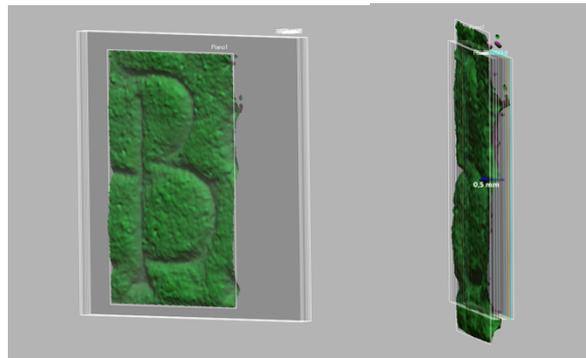


Figure 4.22. Preliminary step for the depth analysis: generation of tangential parallel planes sectioning the last letter of the acronym.

From the analysis, comes up that at 1 mm of depth (Figure 4.23), can be noticed a discontinuity on the bottom part of the letter B, that could indicate an adjustment (blue arc) to a pre-existing letter R, then turned in a B (Figure 4.24).

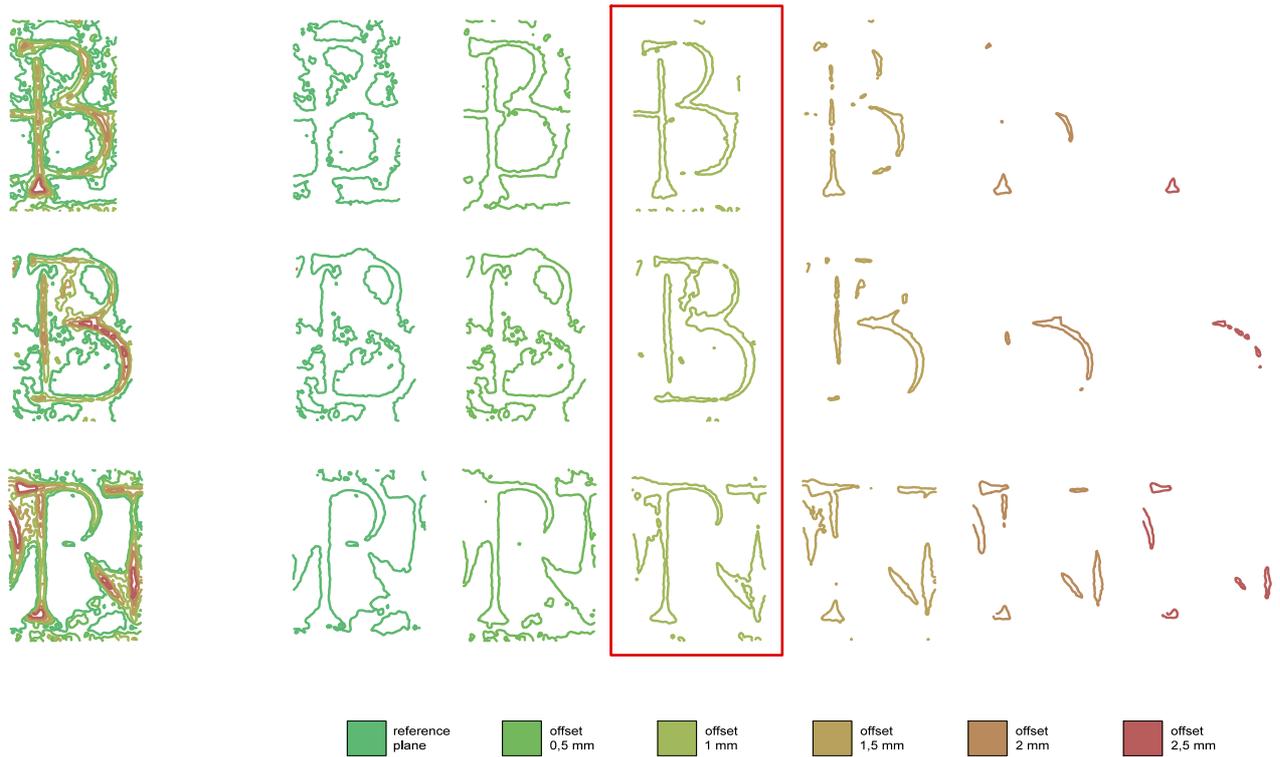


Figure 4.23. Section profiles comparison. From top to bottom: character belonging to the acronym, generic B, generic R: the colored polylines indicating increasing offset distances from the reference surface.

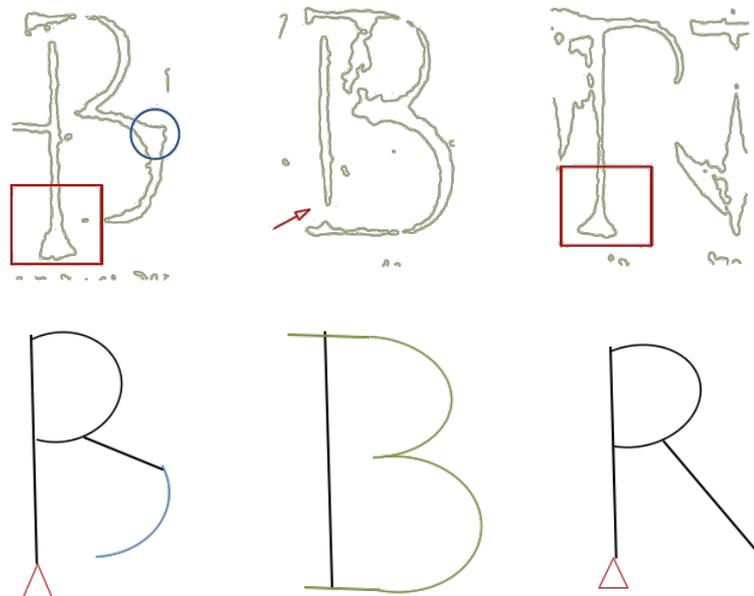


Figure 4.24. Depth analysis at 1,5 mm of depth; from left to right: letter belonging to the acronym, a generic B and a generic R. The discontinuity on the bottom part of the first letter (blue circle) could indicate an adjustment (blue arc) to a pre-existing letter R, then turned in a B. The vertical sign of the debated letter seems to have the same “triangular” shape of the R one (red square).

#### 4.4 Clay tablet and seal of Ancarani Collection, Civic Archaeologic Museum, Bologna

This work focuses on the application of SLS Artec Spider for 3D high-resolution acquisition of small clay artefacts of different shape and size. The surveyed objects belong to the Ancarani Collection of the Civic Archaeological Museum in Bologna. They come from ancient Mesopotamia and bear cuneiform inscriptions, characterized by small signs made up by wedge-shaped elements. This is the reason why this kind of writing, which was used by the ancient inhabitants of Mesopotamia (Sumerians, Babylonians, Assyrians) for more than three thousand years, is called “cuneiform” (from Latin *cuneus*, “wedge”). The objects are a dark-clay tablet and a brick-red seal (Figure 4.25)

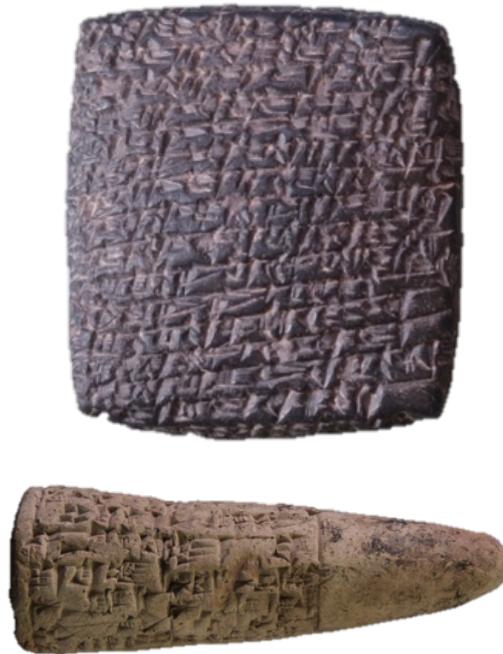


Figure 4.25. Above, dark-clay tablet; below, brick-red seal (images not to scale).

The tablet ( $5,8 \times 5,4 \times 1,8$  cm; museum siglum EG 4013) is a letter in Old Assyrian script and language, sent by a well-known Assyrian merchant, one Puzur-Assur, to two agents of his, and contains a series of instructions of the former for his subordinates concerning various trade and business matters (Kienast, 1960; Michel, 2001). It comes from the ancient city of Kanesh (corresponding to the modern site of Kültepe in central Turkey) and dates to the 19th century BC. As far as the seal is concerned (EG 4200; height: 12,6 cm; diameter at the base: 4,3 cm), we have here an example of royal inscription in Old Babylonian script and Sumerian language. It was commissioned by Ishme-Dagan (ca. 1953-1935 BC), king of Isin (modern Ishan al-Bahriyat in southern Iraq), in order to commemorate his building of the “great wall of Isin” (*bàd gal ì-si-in<sup>ki</sup>-na*). Such objects were usually embedded in the foundation walls of temples or other built structures and served as a sort of time capsule to transmit to the future generations the memory of the king and of his accomplishments.

The first aspect considered in this work is the possibility to acquire surfaces of the finds through the SLS technology and to reconstruct them in high-resolution. Then, are proposed digital procedures for the enhancement of the reconstructed 3D models which, due to the complexity of their surfaces

and their reduced dimensions, are difficult to study. To conclude, a contribution to the automatic extraction and vectorization of the wedges is presented.

#### 4.4.1 3D acquisition and 3D model reconstruction

The survey of both tablet and seal is performed with the hand- guided structured-light scanner, Artec Spider. The survey turns out to be a quite simple and rapid operation because of the small dimensions of the objects and the characteristics of the instrument itself; the guide by hand allows the operator to move around the considered object and to choose the most suitable survey criteria. In both cases the survey approach consists in two successive phases: first the object rotates on a moving support while the scanner is fixed in a position; then, vice versa, the object is fixed in a position and it is the scanner that moves around it. This procedure has been proved to ensure a complete survey without any lack of data and a systematic data acquisition.

Since the scanner is capable to scan the surfaces in *real-time* and to perform the alignment of the frames in an automatic way, the first product is a rough 3D model of the object (i.e. a mesh) that can be visualized and explored by the user immediately after the scanning phase, also permitting to check the quality of data. This first rough model needs a further processing leading to meliorate the mutual alignment of the frames composing the scans and the global alignment between the scans composing the whole object. Moreover, it is characterized by clusters, surface noise, unwanted entities (part of the background involuntarily acquired by the scanner, the support etc.), holes, and unwanted features.

The data processing follows, aiming at the alignment refinement, and mesh correction (manual or automatic). If required, texturing of the final processed model is performed. Problems with texture application have been aroused whit tablet and seal's 3D models, which are related to the high resolution of the model (see paragraph 4.3.2). In this work, two versions of both models have been generated, one textured lower-resolution (1 mm) useful for documentation purpose, and one non-textured high-resolution version (0.3 mm) for the enhancement and wedges extraction.

Usually, the duration of the post-processing highly exceeds the acquisition phase; however, the surveying method described above, together with the dimensions of the objects, ensures an easy and a relatively rapid processing of the data, characterized by a total duration of few hours, editing phase included. The models of the tablet and the cone are shown in Figure 4.26. Table summarizes the model characteristics.



Figure 4.26. Final 3D models. Above: the dark clay tablet textured (left) and non-textured (right). Below: the brick-red seal textured (left) and non-textured (right). Even though the textured models have lower resolution (0,5 mm) compared to the non-textured one (0,3mm), they maintain a satisfying level of detail.

		<i>Face count</i>	<i>Resolution</i>
<b><i>Dark clay tablet</i></b>	<i>Textured</i>	352.643	1 mm
	<i>Non-textured</i>	2.063.641	0.3 mm
<b><i>Brick-red seal</i></b>	<i>Textured</i>	685.926	1 mm
	<i>Non-textured</i>	4.219.644	0.3 mm

Table 4.7. Final models' characteristics.

#### 4.4.2 Digital enhancement

3D models created with the previous mentioned structured light technology become a fundamental starting point for scholars for the philological study of such artefacts. The digital models allow a complete and accurate vision of the entire surface of an inscribed artefact, from every possible point of view, in general and in every detail. However, the visualization limits imposed by their complex geometry, reduced dimensions and small size of wedges, can be overcome only taking advantage of digital operations, which in most of the cases represent the solution to the visualization issues. In fact, the decipherment and interpretation of the inscriptions constituted by very small

engravings that are not easily visible to the naked eye, is a complicated task to accomplish, which however is facilitated by the possibility to apply algorithms, digital filters, and artificial light to the models; these digital solutions provide the enhancement of the cuneiform signs that cover the object surface ensuring easier reading of the text and, therefore, its correct interpretation (Hameeuw and Willems, 2011).

Concerning the digital filtering, mainly two algorithms are used, *Lambertian radiance scaling* (also used for the previous study case) and *Minnaert reflection*, each of them providing different results in terms of surface enhancement. The *Minnaert reflection* algorithm, originally developed for the interpretation of astronomical observations and remote sensing data, starts from the Lambertian BRDF expression and corrects the radiance formulation by introducing a constant  $k$  that weights the incidence and emission angle contributions. This darkens the entity limbs along the edges, providing a 3D model characterized by a high level of surface detail. As Figure 4.27 and 4.28 shows, both algorithms provide good results in terms of wedges enhancement.

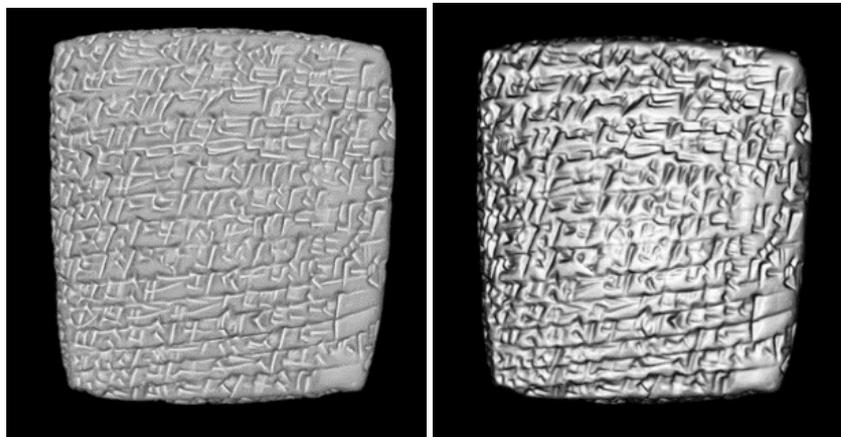


Figure 4.27. Application of radiance scaling (left) and Minnaert reflection (right) to the 3D model of the tablet

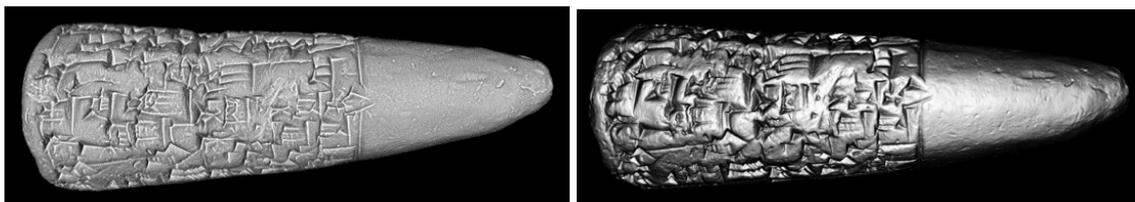


Figure 4.28. Application of radiance scaling (left) and Minnaert reflection (right) to the 3D model of the cone

As in the Pilate's basin, to further improve the visualization of the engraved text, one can take advantage of the application of artificial lights to the digital model (Figure 4.29).

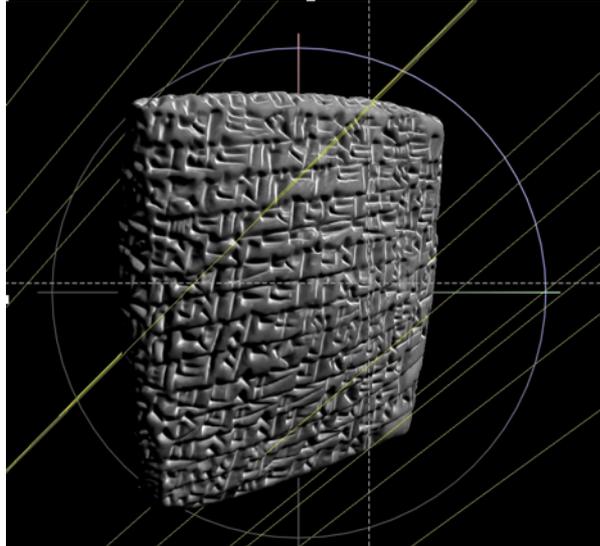


Figure 4.29. Artificial illumination; in yellow, varying direction of the beams of light.

In the case of inscribed artefacts with a conical or cylindrical shape there are specific problems related to the readability of their inscriptions, due to the envelope of their surface (Wagensonner, 2014;). As applied to the circular strip of the basin, an unrolling procedure is necessary for the seal, which can be approximated to a regular cone. The procedure requires the knowledge of geometrical characteristics of the object in order to parametrize the surface. Table 4.8 summarizes the data used for the construction of the regular conic surface: the values of height, radius and aperture angle have been measured digitally on the 3D model of the seal. The algorithm that projects the conic surface on a plane follows. The flat envelope of the cone is shown in Figure 4.30. Besides representing a good visualization improvement and facilitating the reading of the engraved text, the flat reconstruction of the seal constitutes also a digital base for further applications, such as the vectorization of the wedges.

<i>Height</i>	125 mm
<i>Radius</i>	21 mm
<i>Aperture angle</i>	$\approx 30^\circ$

Table 4.8. Regular cone dimensions.

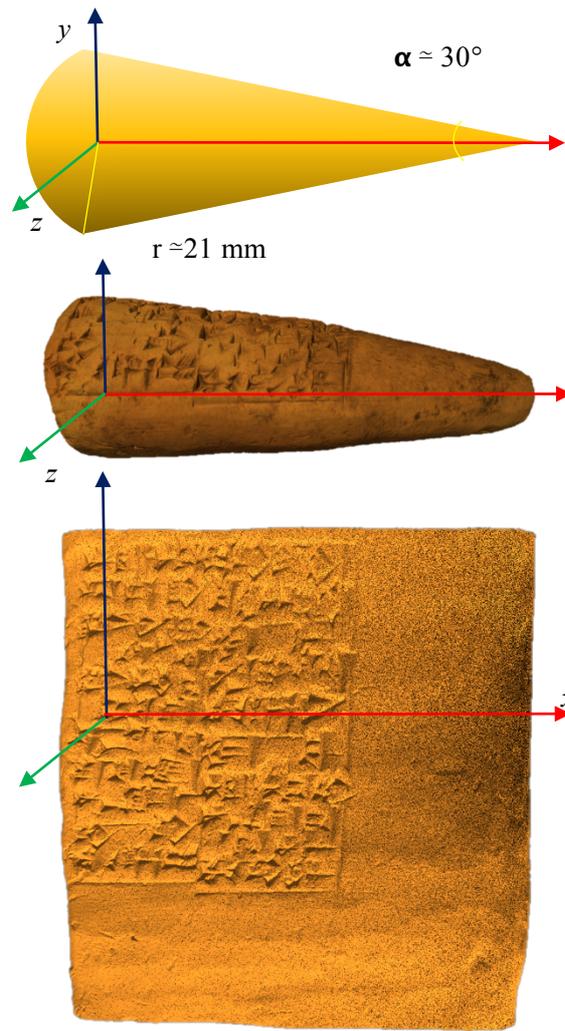


Figure 4.30. Regular cone and seal surface after the unrolling operation.

#### 4.4.3. Vectorization

Simple representation of the artefacts with drawing technique is a frequently used practice especially in the case of epigraphs. Considering the cuneiform scripts, they constitute one of the largest source of ancient history: the clay supports were used for millennia and typically the cuneiform scripts were transcribed by hand on paper with ink. A large number of these documents are stored in a digital library as digital documents in raster graphics, a digital source that is part of a huge project named *Cuneiform Digital Library Initiative (CDLI)* (Figure 4.31).

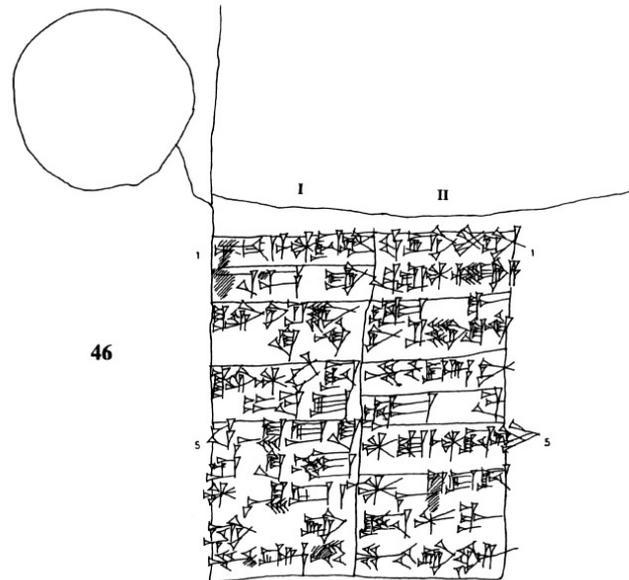


Figure. 4.31. Example of cuneiform script stored in a digital library (<https://cdli.ucla.edu>).

However, one of the main issue related to the hand drawing, resides in the subjectivity of the process, which is influenced also by the experience of the operator. The operations for the wedges representation improvement here proposed aims to lay the basis for the vectorization of the cuneiform inscription to overcome the lacks of the traditional drawing. This issue is nowadays an interest research topic, being a central subject of discussion among philologists. Within this work, the aim is to give a contribution to these studies which still are open questions (Mara and Krömker, 2013; Fisseler et al., 2013; Massa et al., 2016; Rothacker et al., 2015 ); therefore, the vectorization itself fall outside the purpose of this thesis. The process consists in developing a sequence of digital operations capable to give in automatic way, a vector representation of the wedges, which is also unique and objective.

The process starts from the 2D representation of the flat envelope of the seal and applies image processing procedures to further improve the wedges visualization. Here the vectorization steps start, until the extraction of each single wedge. They are performed using the *Potrace* algorithm (Selinger, 2003) which produces a vector outline starting from the decomposition of the bitmap: a graph is constructed following the edges between black and white pixels and then a path is outlined. The approximation of this path results in a polygon, which then is simplified forming an outline, i.e. the vector representation. Depending on the starting image, namely on the degree of wedges enhancement, the vector outline will have different shape and geometry.

Figures 4.33 shows some attempts of vectorization, which are focused on the same group of randomly selected wedges. The parameter that mainly influences the performances of the *Potrace* algorithm are the image processing sequences applied of the unrolled surface of the seal. Three attempts involve the application of *Lit Sphere Radiance Scaling*, *Mean Curvature*, *Grey level occurrence*, and *Lambertian Radiance Scaling*.

*Lit Sphere* method relies on the approximation of the reflected light for a determined material to a reflectivity properties of a lighted sphere of a scale smaller than the scene scale (Sloan et al., 2001). Usually, a different lit sphere is used for convex and concave regions, resulting in further improved legibility of the surface (Figure 4.32). In this case, the result is a well enhanced representation of the

wedges, which appear in high contrast with respect to the base surface. However, since the lit sphere approximates the reflected light on the material, the good characterization of wedges is badly affected by the normals change in correspondence to the engravings, which result described by a color gradient rather than an edged profile. This leads to the “widening” of the wedges tracing and successively, to an ineffective vector representation.

The usage of *Mean Curvature* display mode takes advantage of the computation of the curvature principal directions with the pseudoinverse quadric fitting method. The roundness of the surface is shown using a color range, that in this case relates the red color with negative values of curvature and blue color with positive values. As in the previous case and for similar reasons, the final result after the application of Potrace algorithm is scarce and the wedges are not well described.

Steps ahead have been done with *Grey Level* visualization: the wedges result homogeneously described by uniform distributed levels of gray, even in correspondence of the normals change, but, some irregularities occur due to the roughness of the surface. This effect is evident in Figure 4.33, where in the proximity of three wedges the concave irregularities assumes the same grey level of the wedge, thus invalidating the accuracy of Potrace algorithm. The most satisfying results has been obtained with the Lambertian Radiance scaling shading (Figure 4.34). The wedges profiles are sharp enough to describe faithfully the real engraving shape, allowing the Potrace algorithm to work well until the elaboration of the vector representation.

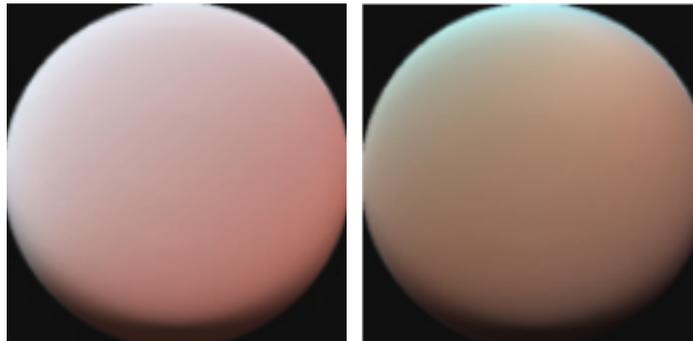


Figure 4.32. Two lit spheres used for *Lit Sphere Radiance Scaling*. On the left, the lit sphere characterizing the convexities, on the right the one describing the concavities behavior.

*Litsphere*

*Mean curvature*

*Grey level*

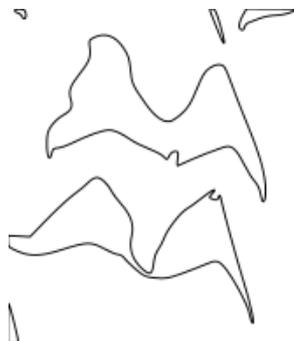
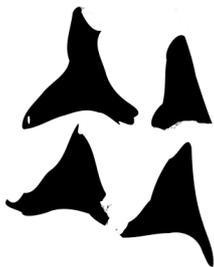
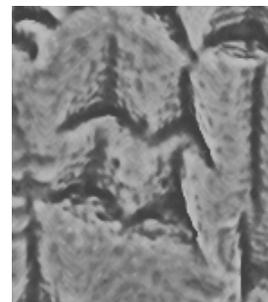
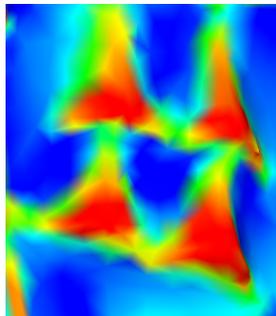
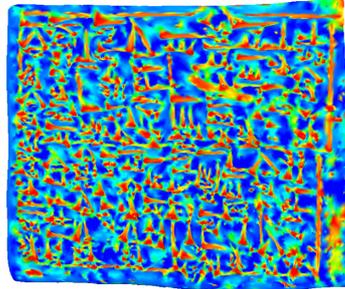
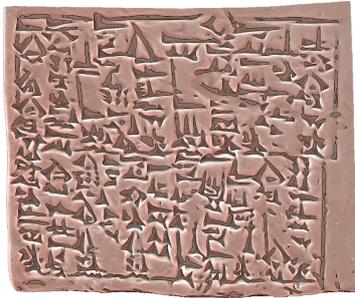


Figure. 4.33. Attempts of vectorization varying the image processing operation for the wedges enhancement. From left to right: lit sphere, mean curvature, and grey level. From top to bottom: the entire unrolled surface after the application of filters, a detail showing a single group of wedges, its binary representation and the vector outline.



Figure 4.34. Vectorization process using inverted Lambertian Radiance scaling: a wedge after image processing procedures, binary representation of the wedge, vector representation, and final product.

The vector representation of wedges is the starting point for creating a drawing of the text much more reliable and objective than a hand-made copy (the so-called “autography”). This operation represents a useful alternative to the 3D model visualization, especially when a simpler representation of the text is required for study, teaching, or publication purposes.

### 4.5 Safe of the Obellio Firmo *domus* in Pompeii

The need for the documentation of the CH is especially felt on the Italian national territory (Cardaci et al., 2019; Tesi et al, 2017) which owns an enormous and valuable patrimony that need to be preserved. Among the numerous sites of historical and archaeological interest, Pompeii is undoubtedly one of the most important, in Italy and in the World. The ancient city, in fact, has been recently the subject of a large project aiming at documenting and mapping the whole archeological site. Taking advantage of the modern surveying techniques, the goal of the “*Grande Progetto Pompei*” has been to build a unique georeferenced information system, at a numerical scale 1:50, which coherently and effectively contains multiple information about the whole structure of the city.

This paragraph describes the experience (Francolini et al., 2020) conducted on a valuable object belonging to Obellio Firmo *domus*, one of the oldest *domus* in Pompeii, located in the northern sector of the ancient city, Regio IX (Figure 4.35, top). The house is the subject of a study conducted by the University of Bologna in agreement with the Archaeological Park of Pompeii (Silani et al., 2017).

The object analyzed is an ancient safe (dimensions 1.50 x 1.20 x 1.00 m), located in the atrium of *domus* (Pesando and Guidobaldi, 2006) next to the main entrance, where the owner used to receive guests or business customers (Figure 4.35, bottom). It represented the high social status of the family living in the *domus* and, in fact, it is located at the entrance of the house as a sign of power and wealth of the family who lives in (Figure 4.36). This type of objects is of particular importance, because there are very few of them, mainly concentrated in this geographical area. The safe of the Obellio Firmo *domus*, fixed on a masonry base, is characterized by unusual decorations that reminds to a knit basket, and round applications which are characteristic elements of this kind of safes having the function of connecting nails between the wooden base and the metal cover.

As can be noticed in Figure 4.36, the safe has currently an irregular shape: during the eruption of Vesuvius in 79 A.D. in fact, the bronze covering and the iron details that compose it, were melted by the heat of the lapilli that reached the urban area of Pompeii and the safe undergone strong deformation. No evidence about the original shape of the safe are known or mentioned in any historical document or book. This work aims to perform a detailed survey of the object in order to create a high-resolution 3D model useful for its documentation, and successively tries to reconstruct the original three-dimensional appearance and shape of the safe basing on the reconstructed 3D model. The choice of the acquisition technique to be applied on the safe went to close-range digital photogrammetry by Multi-View Structure from Motion approach. The reasons are related to the environment in which the case study is located, that did not allow the use of medium size instruments such as those used for 3D scanning (e.g. triangulators or SLS). In fact, as shown in Figure 4.36, to preserve and protect the historical object from atmospheric events and vandalism, a permanent glass case has been installed, whose structure consists of a metal frame. This protection, although indispensable for the good conservation of the safe, was an obstacle to photogrammetric survey, which could not be carried out with a traditional approach but required an unconventional solution.

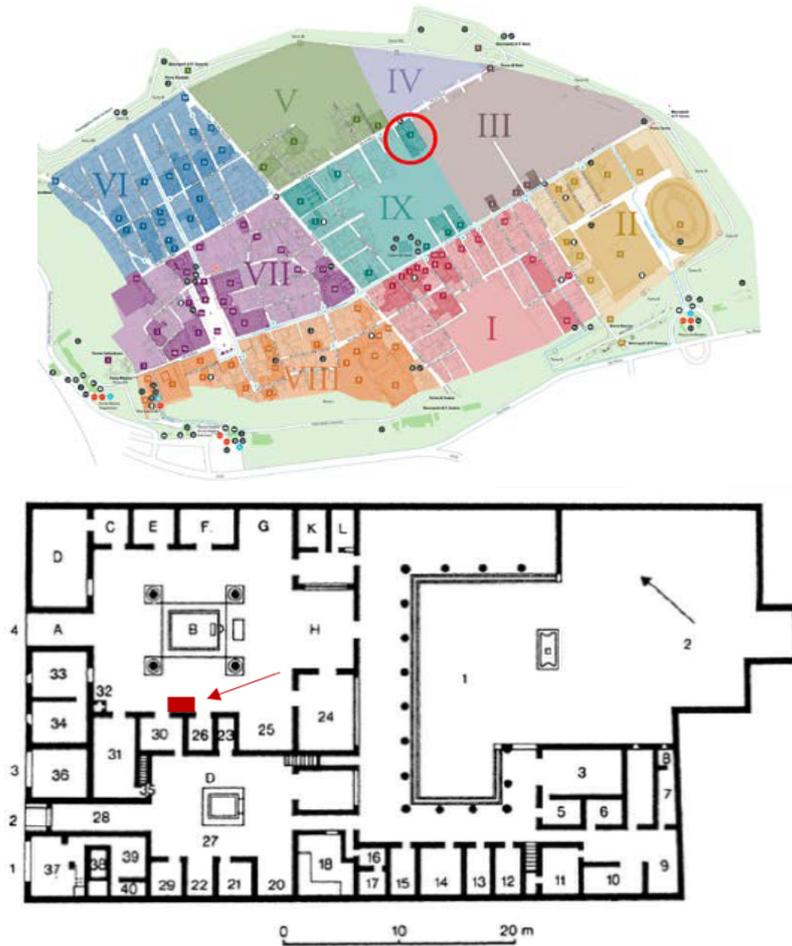


Figure 4.35. Top, Pompeii city map: the red circle indicates the location of the Obellio Firmo domus (<http://pompeiiites.org>). Bottom, Obellio Firmo domus plant: number 4 indicates the main entrance, letter B the main atrium; in red the the location of the safe, in the right side of the atrium (Pesando and Guidobaldi, 2006).



Figure 4.36. The entrance room of the Obellio Firmo domus (left), and the safe (frontal view) (right).

#### 4.5.1 Photogrammetric survey: planning and images acquisition

As explained in Chapter 2, the acquisition planning is a fundamental step for the success of the photogrammetric reconstruction. Considering the survey of the safe, we had to deal with a series of adverse conditions. Besides bad weather conditions and limited time, the major limitation was the environment: the safe in fact is surrounded by a protective glass case installed on a non-removable metal frame and for logistical reasons only the left and right side of the case could be opened at the moment of the survey. The glass case clearly did not allow to perform the photogrammetric acquisition with conventional procedure; so, given these conditions, the survey strategy adopted was to introduce a compact small digital camera (Panasonic TZ60, 1/2.3" CMOS sensor) between the glass surface and the object, moving the camera by a monopod and remotely controlling the acquisition phase. This setup ensured the survey to be carried out safely as the camera was protected from the rain by the glass case in which it was inserted; the distance between the lens and the object was, however, very limited.

It was planned to acquire each side of the safe separately adopting an acquisition configuration that consisted of strips of nadiral and convergent images, acquired in separate steps. Nine coded targets were positioned around the safe at its base, and distances measured between couples of them were used to scale the final model. The team in charge of the survey was composed of two operators, one in charge of holding the camera-monopod system and moving it to make the swipes of images, and the other one to remotely control the camera through a smartphone with a dedicated app (Figure 4.37). As can be observed in Figure 4.36, the object surface is very irregular due to both the folds of the deformed metal covering and the rounded applications on all four sides of the safe: in order to obtain a 3D model capable of accurately describing these peculiarities, a high number of images was acquired. Moreover, the non-conventional remote-controlled camera-monopod system is less immediate and slower than the classic methodology for taking the pictures. These elements, adversely affected the duration of the survey, which took longer than a traditional survey.

598 images were acquired, characterized by 4896 x 3672 pixels, with a 35 mm equivalent focal length of about 25 mm and a pixel size of 1.27 x 1.27  $\mu\text{m}$ . The average distance between the camera and the surface of the safe varies for each side depending on the presence (or absence, in the case of the two lateral sides) of the protective glass. Considering the top and the front of the safe, where the acquisition was hindered by the protective glass, the distance has been quite limited; at an acquisition distance of 20 cm, the resolution on the object is in the order of 0.15 mm/pix and one picture covers a very small area.



Figure 4.37. One operator during the survey holding the camera-monopod system inside the glass case.

#### 4.5.2 Data processing and 3D model generation

The photogrammetric reconstruction of the safe was performed using the Multi-View Structure from Motion approach implemented in the software ContextCapture by Bentley. The overall data processing involved 569 camera stations (almost the entire data set, only few images belonging to the top and front sides were not-aligned and thus rejected) and 9 coded targets for scaling the model. Figure 4.38 shows the very dense photo coverage on the top of the object. Despite being a consumer camera with a wide-angle arrangement of the lens, the distortion grid calculated by the program shows a quite good behavior (Figure 4.39). The processing of the whole model reached good results in terms of accuracy, with a reprojection error (RMS) of 0.87 px. The total number of tie points was around 90000, with a very high average number of points per photo (Figure 4.40). The obtained final 3D model is composed by more than 16 million triangles, and characterized by a very high geometric and radiometric detail; a cleaned model is presented in Figure 4.41.

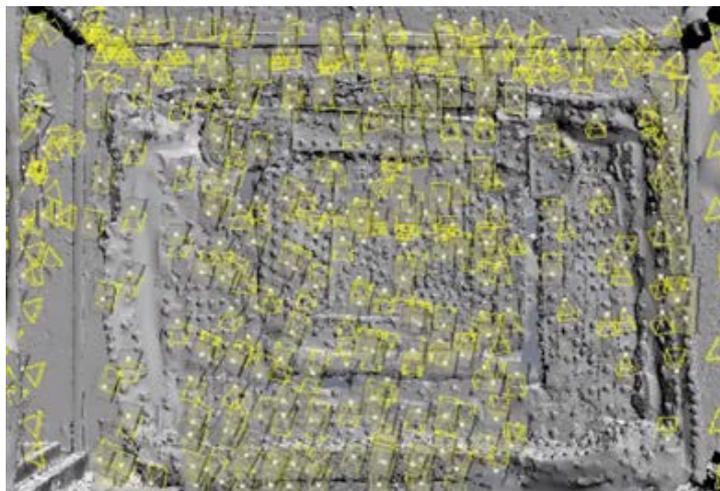


Figure 4.38. Images positions (top view).

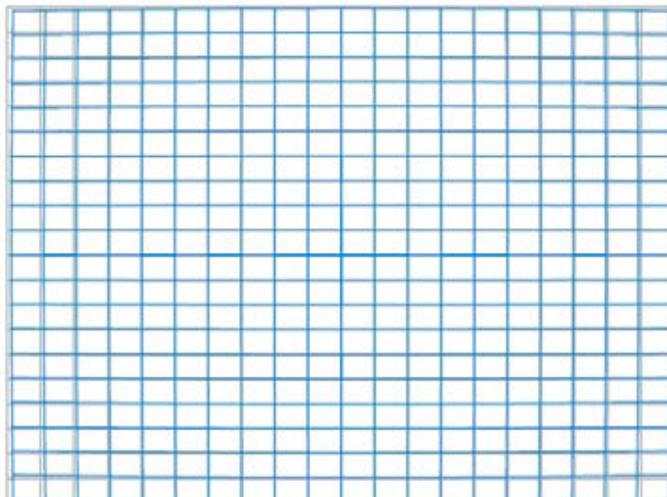


Figure 4.39. The distortion grid.

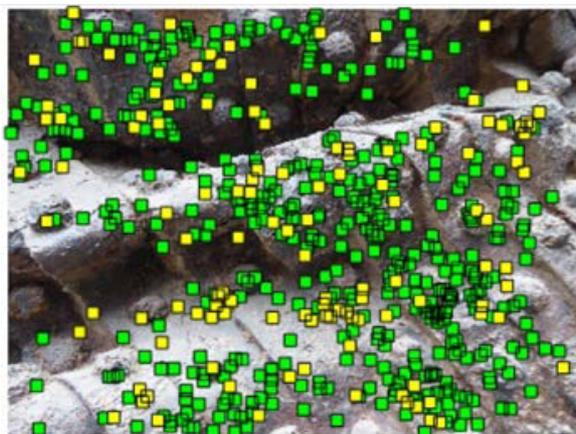


Figure 4.40. Example of an image of the set with the relative tie points.



Figure 4.41. Final 3D model of the safe.

### 4.5.3 3D reconstruction of the original appearance and shape

Since the original appearance and shape of the safe is unknown, the last part of the work aims at reconstructing it, as it was before the Vesuvius eruption. The model has been created on the basis of initial researches on a few other Roman safes in the area (Pernice, 1932; Prisco et al., 2014). Due to the lack of documentation about the safe we are dealing with, it has been assumed the way of opening, as well as some details and the colour. The 3D reconstruction was generated in Rhinoceros; the figure 4.42 shows two sides of the safe.

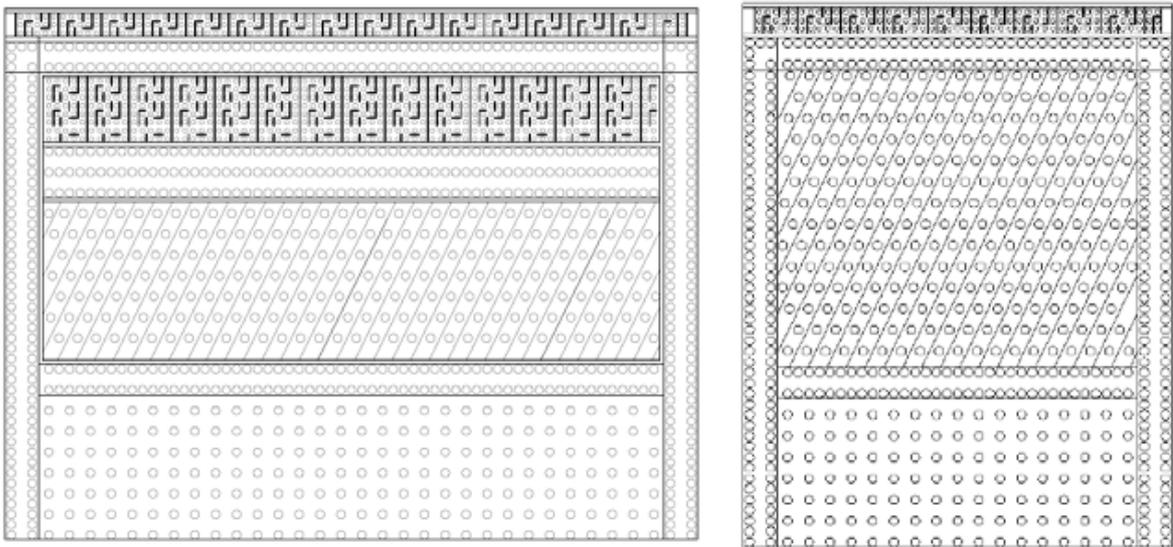


Figure 4.42. Vector drawing of the reconstructed front (left) and side (right) of the safe.

It is important to underline that this reconstruction is only a first attempt. Although made trying to maintain a certain degree of fidelity and rigor, it requires further documentation and knowledge about the ancient Roman safes, which are the starting point for an accurate reconstruction of the details, the decorations and the opening mechanism, features that in the Obellio Firmo safe have been almost lost. Such an analysis could be significant also to support investigations about methods and technologies for producing this type of objects in Roman times.

This work is a further proof of how digital surveying techniques represent a strong support to the CH field. This applies especially when the heritage is particularly valuable and needs a proper documentation for its preservation and valorisation, as the case of the safe of the Obellio Firmo *domus* in Pompeii.

## 5. CONCLUSIONS

The application of 3D high-resolution acquisition techniques has increased during the last decades due to the technological progress related to the sensors devices for the shape acquisition and the development of computer science, which has allowed Geomatics to establish a robust link with the Computer Vision field. Specifically, optical passive and active sensors represent interesting resources in several different fields, from Medicine to Forensic Science, from Cultural Heritage to Geology. In fact, they give the possibility to obtain high-resolution and metrically correct three-dimensional representation of different scale entities without any contact with the surface and in relatively reduced time. Moreover, the need of collecting data for the creation of a digital information database and the accuracy of the data, have contributed to generate great interest about the mentioned surveying techniques coming from applicative fields and academic research.

In this thesis, close-range photogrammetry, ToF laser and structured-light scanners are analyzed starting from the functioning principles of the sensing devices and deepening the issues and the mathematical relations governing the 3D model reconstruction process. Within this work, I wanted to deepen the technical aspects of 3D high-accuracy acquisition not yet sufficiently explored in the literature, highlighting how Geomatics is making a positive contribution to applications generally belonging to other fields and in multidisciplinary contexts.

A first consideration concerns the field of photogrammetry and can be affirmed that there exists the need to review and revisit the traditional photogrammetric methods converting them to Computer Vision field, to which they are now closely related. Nevertheless, the central consideration of this work concerns the analysis of the construction process of three-dimensional models exploiting both close-range photogrammetry and the structured light projection scanners applied on objects of medium and small size. In this context, it is important to underline that, when dealing with applications that require good precision, it is necessary that all the phases that make up the process must be followed carefully, from the acquisition to the processing of data until the creation of the digital model. In particular, we are faced with the need to establish some rules that lead to the creation of a protocol that facilitates the final evaluation of the quality of the work; however, this is not a straightforward task. Therefore, the most convenient strategy we can adopt is to plan and follow carefully all the steps of the three-dimensional reconstruction process.

The phases that mostly require attention and study are the calibration of the instruments, which is crucial for achieving high precision; the acquisition phase that, although it may be ordinary and simple, must be carefully planned modulating it on the type of scenario, environmental conditions, materials and the geometric characteristics of the object to be acquired; and the last phase of interpretation of the results, during which it is not to be excluded the application of techniques not properly distinctive of the context to which we are referring. About this aspect, the last section of the thesis focused on unconventional applications of high-resolution acquisition techniques referring to the field of Cultural Heritage.

From the case studies presented, it emerges that both close-range photogrammetry and SLS bring enormous benefits to the study of objects of medium and small size due to the immediacy of the survey, the ability of 3D models to reach high resolution, their high metric reliability and versatility that makes them real support to the restoration and academic research in the philological and historical field. An innovative aspect this last part of the thesis addresses, is the possibility of

applying algorithms and analysis methodologies coming from other fields, such as Geology and Computer Vision, on the 3D models generated from photogrammetry and SLS. It can be noticed from the results obtained, that these advanced analyses are viable solutions for the quality check of the models as they provide analytical tools that can be easily interpreted and analyzed, thus becoming an integral part of the quality assessment.

## Bibliography

- Acka, D., Remondino F., Novak D., Hanush T., Shorotter G., and Gruen A. (2006): Recording and modeling of cultural heritage objects with coded structured light projection systems. 2nd International Conference on Remote Sensing in Archaeology, Rome, Italy, pp.375-382.
- Aliberti, L., and Iglesias Picazo, P. (2019): Close-range photogrammetry practice: graphic documentation of the interior of the walls of Avila (Spain). ISPRS- International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, XLII-2/W15, pp. 49–53. Doi: 10.5194/isprs-archives-XLII-2-W15-49-2019
- Anwar, H., Din I., and Park K. (2010): Projector Calibration for 3D Scanning Using Virtual Target Images. International Journal of Precision Engineering and Manufacturing, Vol. 13, Issue 1, pp. 125-131. Doi: 10.1007/s12541-012-0017-3
- Artese, S., and Zinno R. (2020): TLS for Dynamic Measurement of the Elastic Line of Bridges. Applied Science, Vol. 19, pp. 1182. Doi: 10.3390/app10031182.
- ASPRS (2013): Accuracy standards for digital geospatial data, Photogrammetric Engineering & Remote Sensing, Vol. 79, issue 12, pp.1073–1085.
- Ballard, D. H., and Brown C.M. (1982): Computer Vision. Prentice Hall.  
ISBN 978-0-13-165316-0.
- Balletti, C., Ballarin M, and Guerra F. (2017): 3D printing: State of the art and future perspectives. Journal of Cultural Heritage Vol. 26, pp. 172-182. ISSN: 1296-2074.  
Doi:10.1016/j.culher.2017.02.010.
- Baltsavias, E. P. (1991): Multiphoto Geometrically Constrained Matching. Ph.D. thesis, Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland, Mitteilungen No. 49, pp. 221.
- Beraldin, J.A., Blais F., Cournoyer L., Godin G., and Rioux M. (2000): Active 3D sensing. Institute of Information Technology, Vol. 10.
- Besl, P.J. (1988): Active, optical range imaging sensors. Machine Vis. Apps. Vol. 1, pp. 127–152. Doi: 10.1007/BF01212277.
- Bitelli, G., Gatta G., Girelli V. A., Vittuari L., and Zanutta A. (2011): Integrated Methodologies for the 3D Survey and the Structural Monitoring of Industrial Archaeology: The Case of the Casalecchio di Reno Sluice, Italy. International journal of geophysics Vol. 2011, pp. 1-8. Doi: 10.1155/2011/874347.

Bitelli G., Girardi F., Girelli V.A. (2014): Digital enhancement of the 3D scan of Suhi I's stele from Karkemish. *Orientalia*, Vol. 83, Issue 2, pp.154-161

Bitelli, G., Balletti C., Brumana R., Barazzetti L., D'Urso M.G., Rinaudo F., and Tucci G. (2017): The GAMHer Research Project for Metric Documentation of Cultural Heritage: Current Developments. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XLII/W1, pp. 239 – 246. Doi:10.5194/isprs-archives-XLII-2-W11-239-2019.

Blais, F. (2004): A review of 20 years of range sensors development. *Journal of Electronic Imaging*, Vol. 13, pp. 231-240. Doi: 0.1117/1.1631921.

Boehler. W, Bordas Vicent M., and Marbs A. (2003): Investigating laser scanner accuracy. *Proceedings of XIX CIPA Symposium, Antalya, Turkey*.

Bonfanti, C., Chiabrando F., and Spanò A. (2010): High accuracy images and range based acquiring for artistic handworks 3D-models. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. 38, pp. 109-114. ISSN 1682-1750.

Borghi, B. (2010): *Viaggio in Terrasanta. La basilica di Santo Stefano in Bologna*. Minerva Edizioni, Bologna.

Brown, D.C. (1971): Close-range camera calibration. *Photogrammetric Engineering*, Vol. 37, Issue 8, pp. 855-866.

Brown, D.C. (1976): The bundle adjustment - progress and prospects, Invited paper of Com. III, XIII. *ISP congress Helsinki*.

Bruno, F., Bruno S., De Sensi G., Luchi M. L., Mancuso S., and Muzzupappa M. (2010): From 3D reconstruction to virtual reality: a complete methodology for digital archaeological exhibition. *Journal of Cultural Heritage* Vol. 11, Issue 1, pp. 42- 49. Doi: 10.1016/j.culher.2009.02.006.

Callieri, M., Cignoni P., Ganovelli F., Impoco G., Montani G., Pingi P., Ponchio F., and Scopigno R. (2004): Visualization and 3D data processing in the David restoration. *IEEE Computer Graphics and Applications*. Vol. 24, Issue 2, pp. 16-21. Doi:10.1109/MCG.2004.1274056.

Cardaci, A., Mirabella-Roberti, G., and Azzola, P. (2019): The Church of Sant'Andrea in Bergamo: an integrated survey for knowledge and conservation. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W15, pp. 239–245. Doi:10.5194/isprs-archives-XLII-2-W15-239-2019.

- Carrero-Pazos, M., and Espinosa-Espinosa D. (2018): Tailoring 3D modelling techniques for epigraphic texts restitution. Case studies in deteriorated roman inscriptions. *Digital Applications in Archaeology and Cultural Heritage*, Vol. 10. ISSN: 2212-0548. Doi:10.1016/J.DAACH.2018.E00079.
- Chen, X., Xi J., Jin Y., and Sun J. (2009): Accurate calibration for a camera-projector measurement system based on structured light projection. *Optic and Lasers in Engineering*. Vol 47, Issue 3-4, pp. 310-319. Doi:10.1016/j.optlaseng.2007.12.001
- Chow, J., Derek L., Lichti D., and William T. (2012): Accuracy assessment of the FARO Focus<sup>3D</sup> and Leica HDS6100 panoramic type terrestrial laser scanners through point-based and plane-based user self-calibration. FIG Working Week 2012. Knowing to manage the territory, protect the environment, evaluate the cultural heritage. Rome, Italy, 6-10 May.
- Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., Ranzuglia, G. (2008): MeshLab: an Open-Source Mesh Processing Tool. Sixth Eurographics Italian Chapter Conference, pp. 129-136.
- Cosentino, S. (2007): Bologna tra la tarda antichità e l'alto medioevo. *Bologna nel medioevo*, pp. 7-104. *Storia di Bologna 2*. Bologna: BUP.
- Dall'Asta, E., Nazarena B., Bigliardi G., Zerbi A., and Roncella R. (2016): Photogrammetric Techniques for Promotion of Archaeological Heritage: The Archaeological Museum of Parma (Italy). *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI/B5, pp. 243-250. Doi: 10.5194/isprsarchives-XLI-B5-243-2016.
- Defrasne, C. (2014): Digital image enhancement for recording rupestrian engravings: applications to an alpine rockshelter. *Journal of Archaeological Science*, Elsevier, Vol. 50, pp 31-38. ISSN 0305-4403. Doi: 10.1016/j.jas.2014.06.010
- Dempski, K., and Viale E., (2005): *Advanced Lighting and Materials with Shaders*. Wordware Publishing Inc. ISBN-13: 978-1556222924.
- Donadio, E., Spanò A., Sambuelli L., and Picchi D. (2018): Three-Dimensional (3D) Modelling and Optimization for Multipurpose Analysis and Representation of Ancient Statues. In *Latest Developments in Reality-Based 3D Surveying and Modelling*; Remondino, F.; Georgopoulos A.; González-Aguilera D.; Agrafiotis P.; Eds.; MDPI: Basel, Switzerland. pp. 95–118. Doi:10.3390/books978-3-03842-685-1/6.
- Dorninger, P., Nothegger C., Pfeifer N., and Molnár G. (2008): On-the-job detection and correction of systematic cyclic distance measurement errors of terrestrial laser scanners. *Journal of Applied Geodesy*. Vol. 2, pp.191–204. ISSN: 1862-9024. Doi: 10.1515/JAG.2008.022.

- Drareni, J., Roy S., and Sturm P. (2009): Geometric video projector auto-calibration. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Miami, pp. 39-46, Doi: 10.1109/CVPRW.2009.5204317.
- Earl, G., Beale, G., Martinez, K., and Pagi, H. (2010): Polynomial texture mapping and related imaging technologies for the recording, analysis and presentation of archaeological materials. ISPRS-International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. 38, pp. 218-223.
- Falcao. G., Hurtos N., Massich J., and Fofi D. (2009): Projector-Camera Calibration Toolbox. <http://code.google.com/p/procamcalib>.
- Fisseler, D., Weichert F., Müller G., and Cammarosano M. (2014): Extending Philological Research with Methods of 3D Computer Graphics Applied to Analysis of Cultural Heritage. 12th EUROGRAPHICS Workshops on Graphics and Cultural Heritage. Doi:10.2312/gch.20141314.
- Fisseler, D., Weichert F., Müller G., and Cammarosano M. (2013): Towards an interactive and automated script feature Analysis of 3D scanned cuneiform tablets. Scientific Computing and Cultural Heritage.
- Forstner, W., and Wrobel B. P. (2016): Photogrammetric Computer Vision. Statistic, Geometry, Orientation and Reconstruction. Springer. ISBN: 978-3-319-11549-8 Doi: 10.1007/978-3-319-11550-4.
- Forsyth, D. A., and Ponce J. (2003): Computer Vision: A Modern Approach. Prentice Hall. ISBN: 978-0-13-085198-7.
- Francolini, C., Marchesi G., and Bitelli G. (2018): High-resolution 3D survey and visualization of Mesopotamian artefacts bearing cuneiform inscriptions. Proceedings IEEE International Conference on Metrology for Archaeology and Cultural Heritage, pp. 178-182. ISBN: 978-1-5386-5275-6.
- Francolini, C., Bitelli G., Borghi B., and Galletti F., (2019): High-resolution 3D surveying in support of Cultural Heritage. Proceedings 2019 IMEKO TC-4 International Conference on Metrology for Archaeology and Cultural Heritage, IMEKO, 2019, pp. 277 – 281.
- Francolini, C., Girelli V.A., and Bitelli G. (2020): 3D image-based surveying of the safe of the Obellio Firmo Domus in Pompeii. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII/B2, pp. 1389 – 1394.
- Fraser, C. (1997): Digital camera self calibration. ISPRS- Journal of Photogrammetry and Remote Sensing. Vol. LII/IV, pp. 149-159. Doi: 10.1016/S0924-2716(97)00005-1.

- Gallinetti, F. (1985): L'iscrizione del catino di Pilato in un contributo inedito di Felice.
- Genta, G., Minetola P., and Barbato G. (2016): Calibration procedure for a laser triangulation scanner with uncertainty evaluation. *Optics and Lasers in Engineering*. Elsevier. Vol. 86, pp. 11-19. Doi: 10.1016/j.optlaseng.2016.05.005.
- Giancola, S., Valenti M., and Sala R. (2018): A Survey on 3D Cameras: Metrological Comparison of Time-of-Flight, Structured-Light and Active Stereoscopy Technologies. ISBN: 978-3-319-91761-0.
- Girelli, V. A., Borgatti L., Dellapasqua M., Mandanici E., Spreafico M. C., Tini M. A., and Bitelli G. (2017): Integration of geomatics techniques for digitizing highly relevant geological and Cultural Heritage sites: the case of San Leo (Italy). *ISPRS- International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W5, pp. 281 – 286. Doi:10.5194/isprs-archives-XLII-2-W5-281-2017.
- Girelli, V.A, Tini M. A., D'Apuzzo M. G., and Bitelli G. (2020): 3D digitization in cultural heritage knowledge and preservation: the case of the Neptune statue in bologna and its archetype. *ISPRS- International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII/B2, pp. 1403 – 1408. Doi: 10.5194/isprs-archives-XLIII-B2-2020-1403-2020.
- Gonzalez, R. C., and Wood R. E. (2018): *Digital image processing*. Pearson International Edition. ISBN-13: 978-0133356724.
- Granshaw, S.I. (1980): Bundle adjustment methods in engineering photogrammetry. The photogrammetric record. Technical Meeting of the Society on 19th February. Vol. 10, issue 56, pp. 181-207. Doi: 10.1111/j.1477-9730.1980.tb00020.x
- Grün, A. (1985): Adaptive least square correlation: a powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography*, Vol. 14, Issue 3, pp. 175-187.
- Grussenmeyer, P., Hanke K., and Streilein A. (2002): Architectural photogrammetry. *Digital Photogrammetry*, M. Kasser and Y. Egels (ed.), Taylor & Francis, pp. 300-339.
- Guarnieri, A., Vettore A., and Remondino F. (2004): Photogrammetry and Ground-based Laser Scanning: Assessment of Metric Accuracy of the 3D Model of Pozzoveggiani Church. *Proceedings of FIG Working week 2004, "The Olympic Spirit in Surveying"*, Athens, Greece.
- Guidi, G., Russo M., and Beraldin J.A. (2010): *Acquisizione 3D e modellazione poligonale*. McGraw-Hill Companies. ISBN-13: 9788838665318.

Guidi, G., Russo M., Magrassi G., and Bordegoni M. (2010): Performance Evaluation of Triangulation Based Range Sensors. *Sensors*. Vol. 10, pp. 7192-7215. ISSN 1424-8220. Doi:10.3390/s100807192.

Guidi, G., and Remondino F. (2012): 3D Modelling from Real Data. *Modeling and Simulation in Engineering*, Catalin Alexandru (Ed.), InTechOpen, Ch. 3. ISBN: 978-953-51-0012-6. Doi: 10.5772/30323.

Guisan, A., Weiss S.B., and Weiss A.D. (1999): GLM versus CCA spatial modeling of plant species distribution. *Kluwer academic publishers. Plant. Ecol.* Vol. 143, pp.107–122.

Hameeuw, H., and Willems G. (2011): New Visualization Techniques for Cuneiform Texts and Sealings. *Akkadica*, Vol. 32, Issue 2, pp. 163–178.

Heller, M., Bauer H. K., Goetze E., Gielisch M., Roth K. E., Drees P., Maier G. S., Dorweiler B., Ghazy A., Neufurth M., Müller W. E., Schröder H. C., Wang X., Vahl C. F., and Al-Nawas B. (2016): Applications of patient-specific 3D printing in medicine. *International Journal of Computerized Dentistry*. Vol. 19, Issue 4, pp. 323-339.

Horn, B. K. P., and Brooks M. J. (Eds.) (1989): *Shape from Shading*. MIT Press, Cambridge, Massachusetts. ISBN: 9780262081832.

Hornberg, A. (2017): *Handbook of machine and computer vision: The Guide for Developers and Users*. Wiley-VCH Verlag GmbH & Co. ISBN: 9783527413409. Doi: 10.1002/9783527413409.

Huang, T.S. (1996): *Computer Vision: Evolution and Promise*. 19th CERN School of Computing. Geneva. Vol. 8, pp. 21-26. ISSN: 0007-8328. ISBN: 9290830956.

Jayanthi, N., and Sreedevi I. (2018): Comparison of image matching techniques, *International Journal of Latest Trends in Engineering and Technology*, Vol.7, Issue 3, pp. 396-400. Doi: 10.21172/1.73.552. e-ISSN:2278-621X.

Jecic, S., and Drvar N. (2003): The assessment of structured light and laser scanning methods in 3d shape measurements. 4th International Congress of Croatian Society of Mechanics. September, 18-20, 2003. Bizovac, Croatia. 237-244.

Kersten, T., Mechelke K., Lindstaedt M., and Strenberg H. (2008): *Geometric Accuracy Investigations of the Latest Terrestrial Laser Scanning Systems. Integrating Generations. FIG Working Week 2008*. Stockholm, Sweden 14-19 June 2008.

- Khoshelham, K., Saadatseresht M., and Gorte B.G. (2012): Influence of tie point distribution on integrated sensor orientation. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* Vol. XXXIV/XXX.
- Kienast B. (1960): *Die altassyrischen Texte des Altorientalischen Seminars der Universität Heidelberg und der Sammlung Erlenmeyer-Basel*. Berlin. De Gruyter, pp. 94-95.
- Kimura, M., Mochimaru M., and Kanade T. (2007): Projector Calibration using Arbitrary Planes and Calibrated Camera. *Proceedings of CVPR, IEEE-Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1–2. Doi: 10.1109/CVPR.2007.383477.
- Kiss, R. (2004): Determination of Drainage Network in Digital Elevation Models, Utilities and Limitations. *Journal of Hungarian Geomathematics*, Vol. 2, pp.16-29.
- Kontogianni, G., Chliverou R., Koutsoudis A., Pavlidis G., and Georgopoulos A. (2017): Enhancing close-up image based 3D digitisation with focus stacking. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XLII-2/W5, pp. 421-425. Doi: 10.5194/isprs-archives-XLII-2-W5-421-2017.
- Kontogianni, G., Lindstaedt M., Kersten T., and Georgopoulos A. (2019): Evaluating the Effect of Using Mirrors in 3D Reconstruction of Small Artefacts. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLII-2/W15, pp. 633-638. Doi: 10.5194/isprs-archives-XLII-2-W15-633-2019.
- Kovacs, L., Zimmermann A. S., Brockmann G., Gühring M., Baurecht H. J., Papadopoulos N., Schwenzer-Zimmerer K. C., Sader R. A., Biemer E., and Zeilhofer H. F. (2005): Three-dimensional recording of the human face with a 3D laser scanner. *Journal of Plastic, Reconstructive and Aesthetic Surgery*. Elsevier. Vol. 59, Issues 1, pp. 1193-1202. Doi: 10.1016/j.bjps.2005.10.025
- Kraus, K. (2007): *Photogrammetry. Geometry from Images and Laser Scans*. De Gruyter, Berlin, (ed. 2) Doi:10.1515/9783110892871.
- Li, Z., Zhu, Q., and Gold, C. (2005): *Digital Terrain Modelling: Principles and Methodology*. CRC Press, Boca Raton, FL, pp. 323.
- Lichti, D., and Harvey B. R. (2002): The Effects of Reflecting Surface Material Properties on Time of Flight Laser Scanner Measurements. *Proceedings of ISPRS Commission IV Symposium 2002 on Geospatial Theory, Processing and Applications, Ottawa, Canada*, Vol. 34.
- Lichti, D., and Jamtsho S. (2006): Angular Resolution of Terrestrial Laser Scanners. *Photogrammetric Record*, Vol. 21, pp. 141-160. Doi: 10.1111/j.1477-9730.2006.00367.x.

- Lowe, D.G. (1999): Object Recognition from Local Scale-Invariant Features. The Proceedings of the Seventh IEEE International Conference on Computer Vision. Vol. 2, pp. 1150-1157. Doi: 10.1109/ICCV.1999.790410.
- Luhmann, T., Robson S., Kyle S., and Bohem J. (2019): Close-range photogrammetry and 3D imaging. De Gruyter. ISBN: 978-3-11-060725-3. Doi: 10.1515/9783110607253.
- Mara, H., and Krömker S. (2013): Vectorization of 3D-Characters by Integral Invariant Filtering of High-Resolution Triangular Meshes. XII International Conference on Document Analysis and Recognition, Washington, DC, pp. 62-66, Doi: 10.1109/ICDAR.2013.21.
- Massa, J., Bogacz B., KrömkeR S., and Mara H. (2016): Cuneiform Detection in Vectorized Raster Images. XXI Computer Vision Winter Workshop. Doi: 10.11588/heidok.00020244.
- Menna, F., Nocerino E., Remondino F., Dellepiane M., Callieri M., and Scopigno R. (2016): 3D Digitization of a Heritage Masterpiece. A critical analysis on quality assessment. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI/B5, pp. 675-683. Doi: 10.5194/isprsarchives-XLI-B5-675-2016.
- Menna, F., Nocerino E., Morabito D., Farella E. M., Perini M., and Remondino F. (2017): An Open Source Low-Cost Automatic System for Image-Based 3D Digitization. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XLII-2/W8, pp. 155-162. Doi: 10.5194/isprs-archives-XLII-2-W8-155-2017.
- Michel, C. (2001): Correspondance des marchands de Kaniš au début du IIe millénaire avant J.-C. Paris: Les éditions du cerf, 2001, pp. 128-129. ISBN: 9782204067003.
- Mills, G., and Fotopoulos G. (2013): On the estimation of geological surface roughness from terrestrial laser scanner point clouds. Geosphere. Vol 9, Issue 5, pp. 1410-1416. Doi: 10.1130/GES00918.1.
- Minto, S., and Remondino F. (2014): Online access and sharing of reality-based 3D models. Sci. Res. Inf. Technol., Vol. 4 pp. 17-28. Doi:10.2423/i22394303v4n2p17.
- Mistretta, F., Satta G., Stochino F., and Vacca G. (2019): Structure from Motion Point Clouds for Structural Monitoring. Remote Sensing. Vol. 11 pp. 1940. Doi: 10.3390/rs11161940.
- Mokarram, M., Roshan, G., and Negahban, S. (2015): Landform classification using topography position index (case study: salt dome of Korsia-Darab plain, Iran). Model. Earth Syst. Environ. Vol. 1, pp. 40. Doi: 10.1007/s40808-015-0055-9.

- Moreno, D., and Taubin G. (2012): Simple, Accurate, and Robust Projector-Camera Calibration. Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission, Zurich, pp. 464-471, DOI: 10.1109/3DIMPVT.2012.77.
- Murtiyoso, A., Grussenmeyer P., Suwardhi D., Fadhila W.A., Permana H.A., and Wicaksono D. (2019): Multi-sensor 3D recording pipeline for the documentation of Javanese temples. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, XLII-2/W15, pp. 829–834, Doi: 10.5194/isprs-archives-XLII-2-W15-829-2019, 2019.
- Neumüller, M., Reichinger A., Rist F., and Kern C. (2014): 3D Printing for Cultural Heritage: Preservation, Accessibility, Research and Education. 3D Research Challenges in Cultural Heritage. Lecture Notes in Computer Science, Vol. 8355. Springer, Berlin, Heidelberg. Doi: 10.1007/978-3-662-44630-0\_9.
- Nocerino, E., Menna F., and Remondino F. (2014): Accuracy of Typical Photogrammetric Networks in Cultural Heritage 3D Modeling Projects. ISPRS- International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-5, Doi:10.5194/isprsarchives-XL-5-465-2014.
- Nouri, M., Farzan A., Akbarzadeh A. R., and Massudi R. (2015): Comparison of clinical bracket point registration with 3D laser scanner and coordinate measuring machine. Dental Press Journal of Orthodontics. Vol. 20 pp. 59–65. Doi: 10.1590/2176-9451.20.1.059-065.oar.
- Oniga, V. E., Breaban A. I., and Statescu F. (2018): Determining the optimum number of Ground Control Points for Obtaining High precision results based on UAS images. Electronic Conference on Remote Sensing. Vol. 2, Issue 7, pp.352. Doi: 0.3390/ecrs-2-05165.
- Oreni, D., Brumana R., Della Torre S., Banfi F., Barazzetti L., and Previtali M. (2014): Survey turned into HBIM: the restoration and the work involved concerning the Basilica di Collemaggio after the earthquake (L'Aquila). ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, II-5, pp. 267–273, Doi:10.5194/isprsannals-II-5-267-2014, 2014.
- Papadaki, A., Agrafiotis P., Georgopoulos A., and Prignitz S. (2018): Accurate 3D scanning of damaged ancient greek inscriptions for revealing weathered letters. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-5/W4, pp. 237-243. Doi: 10.5194/isprsarchives-XL-5-W4-237-2015.
- Pernice, E. (1932): Hellenistische Kunst in Pompeji. Hellenistische Tische, Zisternenmundungen, Beckenuntersätze, Altare und Truhen, pp. 71-94. Berlin-Leipzig, W. de Gruyter & co.
- Perrone, R.V., and Williams J. L. (2019): Dimensional accuracy and repeatability of the NextEngine laser scanner for use in osteology and forensic anthropology. Journal of Archaeological Science. Elsevier. Vol. 25 pp. 308-319. Doi: 10.1016/j.jasrep.2019.04.012.

Pesando, F., and Guidobaldi, M.P. (2006). Gli ozi di Ercole: residenze di lusso a Pompei ed Ercolano, L'Erma di Bretschneider.

Pesci, A., and Teza G. (2008): Terrestrial laser scanner and retro- reflective targets: an experiment for anomalous effects investigation. *International Journal of Remote Sensing*, Vol. 29, Issue 19, pp. 5749-5765, Doi: 10.1080/01431160802108489.

Previtali, M., Barazzetti L., and Scaioni M. (2013): Multi-Step and Multi-Photo matching for accurate 3D reconstruction. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-3/W22. Doi:10.5194/isprsarchives-XXXVIII-3-W22-103-2011.

Prince, S. J.D. (2012): *Computer Vision: Models, Learning, and Inference*. Cambridge University Press, 40 W. 20 St. New York, NY, United States. ISBN: 978-1-107-01179-3.

Prisco, G., Fossà, B., Ferrari, S., Federico, S., Giglio, A., Schneider, K., Scarpitti, P., Priori, G., Talarico, F., and Villa, I.M. (2014): La cassaforte della casa dei Vettii a Pompei, *Bollettino ICR* n.28, Istituto Superiore per la Conservazione ed il Restauro, pp. 68-86. ISCR-Nardini Editore.

Quattrini, R., Malinverni E. S., Clini P., and Nespeca R. (2015): From TLS to HBIM: high quality semantically-aware 3D modeling of complex architecture. *ISPRS- International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-5/W4, pp. 267-374. Doi: 10.5194/isprsarchives-XL-5-W4-367-2015.

Remondino F., and Fraser C. (2006): Digital camera calibration methods: consideration and comparisons. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVI/5, pp. 266-272.

Remondino, F. (2011): Heritage Recording and 3D Modeling with Photogrammetry and 3D Scanning. *Remote Sensing*. Vol. 3, Issue 6, pp. 1104-1138. Doi: 10.3390/rs3061104.

Rioux, M. (1994): Digital 3-D imaging: theory and applications. *Proceedings SPIE 2350, Videometrics III*, Vol. 2350, pp. 2-15. Doi: 10.1117/12.189122.

Rothacker, L., Fisseler D., Müller G. G. W., Weichert F., and Fink G. A. (2015): Retrieving Cuneiform Structures in a Segmentation-Free Word Spotting Framework. *Proceedings of 3rd International Workshop on Historical Document Imaging and Processing*. ISBN: 9781450336024. Doi: 10.1145/2809544.2809562.

Sansoni, G., Trebeschi M., and Docchio F. (2009): State-of-The-Art and Application of 3D Imaging Sensors in Industry, Cultural Heritage, Medicine and Criminal Investigation. *Sensors*. Vol. 9, Issue 1, pp. 568-601. Doi: 10.3390/s90100568.

Salvi, J., Pàges J., and Batlle J. (2003): Pattern codification strategies in structured light systems. *Pattern Recognition*. Elsevier. Vol. 37, pp 827-849. Doi: 10.1016/j.patcog.2003.10.002

Selinger, P. (2003): Potrace: a polygon-based tracing algorithm. <http://potrace.sourceforge.net/potrace.pdf>

Shan, J., and Toth, C.K. (Eds.) (2017): *Topographic laser ranging and scanning: principles and processing* (2<sup>nd</sup> ed.). CRC Press. Doi: 10.1201/9781315154381.

Shannon, C. E. (1948): A Mathematical Theory of Communication. *Bell System Technical Journal*, Vol. 27, Issue 3, pp. 379–423. Doi: 10.1002/j.1538-7305.1948.tb01338.x.

Shilts, S.B., Wrmlnder S., Louise M. F., Miller K., and Walker P.L. (2010): Variation in the measurement of cranial volume and surface area using 3D laser scanning technology. *Forensic Sci*, July 2010, Vol. 55, No. 4. Doi: 10.1111/j.1556-4029.2010.01380.x.

Silani, M., Giorgi, E., Boschi, F., Bitelli, G., and Martellone, A. (2017): Seeing into the past: integrating 3D documentation and non-invasive prospecting methods for the analysis, understanding and reconstruction of the ancient Pompeii. The case of the House of Obellio Firmo (IX, 14). *Archeologia e Calcolatori*, Volume 28, Issue 2, pp. 361-367.

Sloan P., Martin W., Gooch A., and Gooch B. (2001): The Lit Sphere: A Model for Capturing NPR Shading from Art. *Proceedings of the Graphics Interface 2001 Conference*, June 7-9 2001, Ottawa, Ontario, Canada, pp. 143-150.

Sobel, I. (2014): An Isotropic 3x3 Image Gradient Operator. Presentation at Stanford A.I. Project 1968.

Soudarissanane, S., Lindenbergh R., Menenti M, and Teunissen P. (2011): Scanning geometry: Influencing factor on the quality of terrestrial laser scanning points. *ISPRS- International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* Vol. 66, Issue 4, pp- 389–399. Doi:10.1016/j.isprsjprs.2011.01.005.

Spreafico, M. C., Francioni M., Cervi F., Stead D., Bitelli G., Ghirotti M., Girelli V. A., Lucente C. C., Tini M. A., and Borgatti L. (2016): Back Analysis of the 2014 San Leo Landslide Using Combined Terrestrial Laser Scanning and 3D Distinct Element Modelling. *Rock Mechanics and Rock Engineering*. Vol. 49, pp. 2235 – 2251. Doi: 10.1007/s00603-015-0763-5.

Stjepandić, J., Wognum N., and Verhagen W. (2015): *Concurrent Engineering in the 21st Century. Foundations, Developments and Challenges*. [Springer International Publishing](https://www.springer.com/9783319137766). ISBN: 978-3-319-13776-6.

Strahler, A. N. (1957): Quantitative analysis of watershed geomorphology. *Eos Trans. AGU*. Vol. 38, Issue 6, pp. 913– 920, Doi:10.1029/TR038i006p00913.

Sturm, P., and Maybank S. (1999): On plane-based camera calibration: A general algorithm, singularities, applications. *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Fort Collins, CO, USA, 1999, Vol. 1, pp. 432-437, Doi: 10.1109/CVPR.1999.786974.

Tan, K., Zhang W., Shen F., and Cheng X. (2018): Investigation of TLS Intensity Data and Distance Measurement Error from Target Specular Reflections. *Remote Sensing*. Vol. 10, Issue 7, pp. 1077. Doi: 0.3390/rs10071077.

Tesi, V., Tucci, G., Bonora, V., Fiorini, L., and Conti, A. (2017): Laser scanning and modelling of barely visible features: the survey of the Grotto of the Animals at the Villa of Castello (Florence), *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-5/W1, pp. 343–349, Doi: 10.5194/isprs-archives-XLII-5-W1-343-2017.

Thali, M. J., Braun M., and Dirnhofer R. (2003): Optical 3D surface digitizing in forensic medicine: 3D documentation of skin and bone injuries. *Forensic Science International*. Vol.137, Issue 2-3, pp. 203-208. Doi:10.1016/j.forsciint.2003.07.009.

Triggs B., McLauchlan P.F., Hartley R.I., Fitzgibbon A.W. (2000): Bundle Adjustment — A Modern Synthesis. In: Triggs B., Zisserman A., Szeliski R. (eds) *Vision Algorithms: Theory and Practice*. IWVA 1999. *Lecture Notes in Computer Science*, vol 1883. Springer, Berlin, Heidelberg. Doi: 10.1007/3-540-44480-7\_21.

Ullman S. (1979): The interpretation of Structure from Motion. *Proceedings of Royal Society of London*. Vol. 203, Issue 1153, pp. 405–426. Doi:10.1098/rspb.1979.0006.

Valkenburg, R. J., and McIvor A. M. (1998): Accurate 3D measurement using a structured light system. *Image Vision and Computing*. Elsevier. Vol. 16, pp. 99-110. Doi: [https://doi.org/10.1016/S0262-8856\(97\)00053-X](https://doi.org/10.1016/S0262-8856(97)00053-X).

Vergne, R., Pacanowski R., Barla P., Granier X., and Schlick C., (2010): Radiance Scaling for versatile surface enhancement. *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, pp. 143-150. Doi: 10.1145/1730804.1730827.

Waldhäusl, P., Ougleby C. (1994): 3x3 Rules for Simple Photogrammetric Documentation of Architecture. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXX5, pp. 426-429.

Wagensonner, K. (2014): Digitizing in the round. *Cuneiform Digital Library Notes*. Vol. 8.

- Weide, D. L., and Webster, G. D. (1967): Ammonium Chloride Powder Used in the Photography of Artifacts. *American Antiquity*, Society for American Archaeology Vol. 32, Issue 1, pp. 104-105. Doi: 10.2307/278783.
- Wenfei, X., Zhengtao S., and Dongsheng L. (2017): Comparisons of feature extraction algorithm based on unmanned aerial vehicle image. *Open Phys.* 2017; Vol. 15, pp.472–478. De Gruyter Open. DOI 10.1515/phys-2017-0053.
- Wenzel, K., Rothermel M., Fritsch D., and Haala N. (2013): Image acquisition and Model Selection for Multi-View Stereo. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-5/W1. Doi: 10.5194/isprsarchives-XL-5-W1-251-2013.
- Xi, S. Z., and Donsheng Li (2017): Comparison of feature extraction algorithm based on unmanned aerial vehicle image. *De Gruyter, Open Physics*, Vol. 15, Issue 1, pp. 472–478. eISSN: 2391-5471 Doi: 10.1515/phys-2017-0053.
- Xu, X., Yang H., and Neumann I. (2018): A feature extraction method for deformation analysis of large-scale composite structures based on TLS measurement. *Composite Structures*. Elsevier. Vol. 184 pp. 591-596. Doi: 10.1016/j.compstruct.2017.09.087
- Zámečníková, M., Wieser A., Woschitz H., and Ressel C. (2014): Influence of surface reflectivity on reflectorless electronic distance measurement and terrestrial laser scanning. *Journal of Applied Geodesy*. Vol. 8, Issue 4, pp. 311-325. Doi: 10.1515/jag-2014-0016.
- Zappa, E., Busca G., and Sala P. (2010): Innovative calibration technique for fringe projection based 3D scanner. *Optics and Lasers in Engineering*. Elsevier. Vol. 49, Issue 3, pp. 331-340. Doi:10.1016/j.optlaseng.2010.10.007.
- Zhang, R., Tsai P. S., Cryer J. E., and Shah M. (1999): Shape from Shading: A Survey. *Pattern Analysis and Machine Intelligence*, IEEE Transactions, Vol. 21. Doi: 10.1109/34.784284.
- Zhang, Z. (2000): A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence*, IEEE Transactions, Vol. 22, Issue11, pp.1330–1334. Doi: 10.1109/34.888718.
- Zhang, Z. (2004): Camera Calibration with One-Dimensional Objects. *Pattern Analysis and Machine Intelligence*, IEEE Transactions, Vol. 26, Issue 7, pp.892–899. Doi:10.1109/TPAMI.2004.21.
- Zhilin, L., Zu Q., and Gold C. (2005): *Digital Terrain Modeling. Principles and methodology*. Boca Raton, CRC Press. Doi:10.1201/9780203357132.
- Ziou, D., and Tabbone S. (1998): Edge detection techniques: An overview. *International Journal of Pattern Recognition and Image Analysis*, Vol. 8, Issue 4, pp. 537–559.